

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2002, 16:42:00 ; Search time 1894 seconds
(without alignments)
13623.242 Million cell updates/sec

Title: US-09-598-443-1
1233
Sequence: 1 atgcgaagtgctgtgatag.....tgtccaagatgatatgtag 1233

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

St. ed: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*

1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vl:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*
29: em.vl:*
30: em.htg.hum:*
31: em.htg.inv:*
32: em.htg.other:*
33: em.htg.inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description
------------	-------------	--------	----	-------------

1	1233	100.0	1627	9	BC003591	BC003591 Homo sapi
2	1233	100.0	1695	9	AK025099	AK025099 Homo sapi
3	1231.4	99.9	1659	6	AX058610	AX058610 Sequence
4	1071.8	86.9	1649	6	AX201368	AX201368 Sequence
5	860.2	69.8	1568	10	BC010806	BC010806 Mus muscu
6	856.6	69.5	1421	10	AF239957	AF239957 Mus muscu
7	826	67.0	1540	10	AF113795	AF113795 Mus muscu
8	141.6	11.5	61326	2	AC108908	AC108908 Mus muscu
9	122	9.9	324	11	G38549	G38549 SHGC-64262
10	113	9.2	62831	2	AC109272	AC109272 Mus muscu
11	73.6	6.0	61326	2	AC108908	AC108908 Mus muscu
12	68	5.5	125020	9	AF429315	AF429315 Homo sapi
13	65.4	5.3	110000	2	AC091454_2	Continuation (3 of
14	64	5.2	145936	2	AC105364	AC105364 Oryza sat
15	63	5.1	207683	2	AC098712	AC098712 Mus muscu
16	62.6	5.1	125020	2	AF429315	AF429315 Homo sapi
17	60.2	4.9	318930	2	AC073495	AC073495 Mus muscu
18	59	4.8	241432	10	AL589661	AL589661 Mouse DNA
19	58.6	4.8	120598	2	OSJN00104	AL606653 Oryza sat
20	58.6	4.8	142799	2	OSJN00149	AL663017 Oryza sat
21	57	4.6	451	10	M581R3E2	M10668 Mouse DNA w
22	57	4.6	156654	8	AC074283	AC074283 Oryza sat
23	56.8	4.6	110000	2	LMFLCHR36_31	Continuation (32 o
24	56.8	4.6	175296	2	AC106674	AC106674 Rattus no
25	56.6	4.6	110000	2	LMFLCHR32_06	Continuation (7 of
26	56.4	4.6	1150	14	HS4ULIR3	J02079 epstein-bar
27	56.4	4.6	1818	9	GORINOLIB	M23604 Gorilla gor
28	56.4	4.6	1926	6	AX107940	AX107940 Sequence
29	56.4	4.6	2580	6	AR108994	AR108994 Sequence
30	56.4	4.6	5452	6	AR083151	AR083151 Sequence
31	56.4	4.6	5452	12	U02454	U02454 Cloning vec
32	56.4	4.6	9600	6	A92665	A92665 Sequence 1
33	56.4	4.6	9600	6	IR5041	IR5041 Sequence 15
34	56.4	4.6	10596	6	I25041	I25041 Sequence 15
35	56.4	4.6	10596	6	I30503	I30503 Sequence 15
36	56.4	4.6	10737	12	XXU02428	U02428 Cloning vec
37	56.4	4.6	10850	12	U02455	U02455 Cloning vec
38	56.4	4.6	172281	14	EBV	V01555 Epstein-Bar
39	56.4	4.6	184113	14	HS4958RAJ	M080517 Epstein-Bar
40	55.8	4.5	38543	1	SCBAC16H6	AL596162 Streptomy
41	55.8	4.5	192675	10	AL589767	AL589767 Mouse DNA
42	55.6	4.5	229896	14	AF232689	AF232689 Rat cytom
43	55.2	4.5	1458	8	AF387790	AF387790 Sorghum b
44	55.2	4.5	138070	2	AC109783	AC109783 Mus muscu
45	55.2	4.5	224061	2	AL606909	AL606909 Mus muscu

ALIGNMENTS

RESULT 1 BC003591 1627 bp mRNA linear PRI 12-JUL-2001
LOCUS Homo sapiens, single Ig IL-1R-related molecule, clone MGC:4338
DEFINITION IMAGE:2821373, mRNA, complete cds.

ACCESSION BC003591
VERSION BC003591.1 GI:13097794

KEYWORDS MGC.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 1627)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (20-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DRP

CDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (L1NL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada

Steven Jones, Jennifer Asano, Ian Bosdel, Aaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzyżniński, Reta Kutscher, Oliver Lee, Soo Leela, Victor Ling, Carrie Mathewson, Candice McLeary, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhū, Parvaneh Speedi, Jacqueline Schein, Duane Smallus, Michael Smith, Loraine Spence, Jeff Stott, Michael Thorne, Miranda Tsai, George Yang, Scott Zuydamduin, Naastja van den Bosch, Jill Vardy, Marco Marra.

clone distribution. MCC clone distribution information can be found through the I.M.A.G.E. Consortium/INL at: <http://Image.INL.gov>
Series: IRML Plate: 9 Row: n Column: 3
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA g.i.: 10437546.

FEATURES

Source

```
1. 1637 "Organism:"Homo sapiens"  
db_xref="locusid:59307"  
db_xref="taxon:9606"  
clone:MGC:4338 IMAGE:2821373"  
tissue_type:"lung, small cell carcinoma"  
clone_id:"NH_MGC_7"  
lab_host:"DHIOB-R"  
note:"Vector: pOTB"  
157. 1389
```

CDS

```

/codon_start=1
/product="single Ig IL-1R-related molecule"
/protein_id="AAH03591.1"
/db_xref="GI:13097795"
/translation="MPCVCDRAPDPLSPEDVLRPALGSSVALNCTAMVYSGHSL
PSVWLKDGDLGIGGYSLHXYSWKANKSEIVYSAGVSTETWYGGSEION
ISFSETLORAGSPSHVAATVLAIVLLALALYYKRLNVLVODAGVEIN
DKFTLIVLSDVSDCPEDRKFNVLTLKRLERRGYFLPDDNDILPRAPSDAILVNL
SCKRLIVLSDVSLRANMCSHFREGLCGLDLETRFPIFFEGEORDRPHAPLRL
RHLRLVLLLMRPGSVTPSSDEPMKVEYOLALPKYRVREVEDTQLODDKPMILL
RVRVEGRALDSEVDDPDDEGLGVRGVPGEBSAPRHTGSVLSGRSSSEVDSLDGS
RVYSARTGYPLVYSKDM"
BASE COUNT      325 a      292 t
ORIGIN          524 c      486 g

```

Query Match	100.0%	Score 1233;	DB 9;	Length 1627;
Best Local Similarity	100.0%	Pred. No. 1.1e-178;		
Matches 1233;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	ATGCGAGGATGTCGATAGGAGCCCTGACTTCCTCCGCCGTCGTAAGACAGGTCG	60
D	157	ATGCGAGGATGTCGATAGGAGCCCTGACTTCCTCCGCCGTCGTAAGACAGGTCG	216
QY	61	AGGCGTGGCTTGGGACGCTCAGTGGCTCTGAACGACGCTTGGGATGTCCTGGGCC	120
D	217	AGGCGTGGCTTGGGACGCTCAGTGGCTCTGAACGACGCTTGGGATGTCCTGGGCC	276
QY	121	CACGTGCTCCCTGCTTCACATCGAGGCTGAAAGAGGCTTCGATGGGAATTGGGGGC	180
D	277	CACGTGCTCCCTGCTTCACATCGAGGCTGAAAGAGGCTTCGATGGGAATTGGGGGC	336
QY	181	CACTAACAGCTTCACAGATGACTCCCTGGGTCACAGGCCAACTCTGAGAGTGGTCTGTCC	240
D	337	CACTAACAGCTTCACAGATGACTCCCTGGGTCACAGGCCAACTCTGAGAGTGGTCTGTCC	396
QY	241	AGTGTCTGGGGGTCAACGTGACGACCTGAAGTCTATGGGGCTTCACCTGCTTCATC	300
D	397	AGTGTCTGGGGGTCAACGTGACGACCTGAAGTCTATGGGGCTTCACCTGCTTCATC	456
QY	301	CAGAACATCAGCTTCTCCTCCTTCCTACTTTAGAGACCTGGCCCTACAGGCCAGTGGCT	360
D	457	CAGAACATCAGCTTCTCCTCCTTCCTACTTTAGAGACCTGGCCCTACAGGCCAGTGGCT	516

[illegible]

AUTHORS Kawabata,A., Hiki,J.I., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
TITLE MEDO human cDNA sequencing project
JOURNAL Unpublished (2000)
REFERENCE 2 (bases 1 to 1695)
AUTHORS Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T., Shibahara,T., Tanaka,T. and Nakamura,Y.
TITLE Direct Submission
JOURNAL Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center, Shirokane-dai, 4-6-1, Minato-Ku, TOKYO 108-8639, Japan (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)
COMMENT MEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan: cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).

FEATURES
source location/Qualifiers

1..1695
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="COL4458"
/tissue_type="colon"
/clone_id="COL"
/note="cloning vector pME18SFL3"
227..1459
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB15066.1"
/db_xref="GI:10437547"
/translation="MPGYCDRAPDPLSPEDVLRPALGSSVALNCTAMVSGPRCSL
PSVWLKDLGPLEIGTSHLHYSWKANLVSIVLVGNVSTEVQADVECTI
ISFSEFTLQRAQPTSHVAIVLALSLVLLALVLAALVYKRLNVLWYDAAGEVIN
DKLTDAYVSDCPEDRKFNFIILKPOLERRRGKILFDRLPLRAEPSADLVLN
SCRRLIYVLSDAFLSRAMCSHFREGCLRLLELRPIETFEQORBDPAIPALRL
ROHRLVLLLRPGSVTPSSPDMEVOLAIPRKVRYRVEDPOTLODDKDPMLIL
RGRVPEGRALDSEVDPDPEGLDVGKVPFPGERSAPPHISGVLSGRSSEVDVSDLS
RNYSAKRTFYCLVSKDM"

BASE COUNT 329 a 556 c 510 g 300 t
ORIGIN

Query Match 100.0% Score 1233; DB 9; Length 1695;
Best Local Similarity 100.0% Pred. No. 1,1e-178;
Matches 1233; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGCCAGTGTCTGTGATAGGGCCCTGACTTCCTCCCGCTGTAAGCAGGTGTG 60
DB 227 ATGCCAGGTCTGTGATAGGGCCCTGACTTCCTCCCGCTGTAAGCAGGTGTG 286
OY 61 AGGCTGCTTGGGAGCTCAAGTGGCTGAACTGCACGGCTTGGTACTCTTGGGCC 120
DB 287 AGGCTGCTTGGGAGCTCAAGTGGCTGAACTGCACGGCTTGGTACTCTTGGGCC 346
OY 121 CACTGCTCCCTGCTCACTCCAGTGGCTGAAGAAGGGCTTCCATTGGGAATTGGGGCC 180
DB 347 CACTGCTCCCTGCTCACTCCAGTGGCTGAAGAAGGGCTTCCATTGGGAATTGGGGCC 406
OY 181 CACTACAGCTCCAGAGTACTCTGGTCAAGGCCAAGCTGTCAAGAGTGTGTCTC 240
DB 407 CACTACAGCTCCAGAGTACTCTGGTCAAGGCCAAGCTGTCAAGAGTGTGTCTC 466
OY 241 AGTGTCTGGGGGTCAAGTCAACAGCTGAAGTCTATGGGGCTTCACTGTCTCAATC 300
DB 467 AGTGTCTGGGGGTCAAGTCAACAGCTGAAGTCTATGGGGCTTCACTGTCTCAATC 526
OY 301 CAGACATAGCTTCTCTCACTCTTCAAGAGAGTGGGGCTTCAAGCCAGTGGGT 360
DB 527 CAGACATAGCTTCTCTCACTCTTCAAGAGAGTGGGGCTTCAAGCCAGTGGGT 586

OY 361 GGGTGTGAGCCCTCCCTCTGTGCTGTGAGCCCTGTGCTGTGAGCCCTGTCTATGTC 420
DB 587 GGGTGTGAGCCCTCCCTCTGTGCTGTGAGCCCTGTGCTGTGAGCCCTGTCTATGTC 646
OY 421 AAGTGTGCTCAAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 480
DB 647 AAGTGTGCTCAAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 706
OY 481 GAGGGAGCTCTACAGAGCTTACAGCTTCTACAGAGCTTCTACAGAGCTTCTACAG 540
DB 707 GAGGGAGCTCTACAGAGCTTACAGAGCTTCTACAGAGCTTCTACAGAGCTTCTACAG 766
OY 541 GTGAATTCATCTTAAAGCCAGCTGAGAGCGGCTGAGAGCTTCTTCTGAGC 600
DB 767 GTGAATTCATCTTAAAGCCAGCTGAGAGCGGCTGAGAGCTTCTTCTGAGC 826
OY 601 GACCGGAGCTCTGTGCGGCGCTGAGAGCTTCTGCGGAGCTTGTGTACCTGAGCGC 660
DB 827 GACCGGAGCTCTGTGCGGCGCTGAGAGCTTCTGCGGAGCTTGTGTACCTGAGCGC 886
OY 661 TGCAGAGCTCATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 720
DB 887 TGCAGAGCTCATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 946
OY 721 AGCTTCCGGGAGGCGCTGTGCGGCTGTGAGCTTACCCGAGACCATCTTCATCAC 780
DB 947 AGCTTCCGGGAGGCGCTGTGCGGCTGTGAGCTTACCCGAGACCATCTTCATCAC 1006
OY 781 TTGAGAGGCGCAGAGCGGCAAGCCCGGACCGGCGCTTCCGCTGTGTGCGCAGACCGC 840
DB 1007 TTGAGAGGCGCAGAGCGGCAAGCCCGGACCGGCGCTTCCGCTGTGTGCGCAGACCGC 1066
OY 841 CACCTGTGACCTGTGCTGTGCTGTGAGAGCGCGGCTGCTGACTTCTTCCGATTTTGG 900
DB 1067 CACCTGTGACCTGTGCTGTGCTGTGAGAGCGCGGCTGCTGACTTCTTCCGATTTTGG 1126
OY 901 AAGAAGTCACTGCGCTGCGCGGAGAGTGCAGTACAGGCGCGGTGGAAGAGACCCC 960
DB 1127 AAGAAGTCACTGCGCTGCGCGGAGAGTGCAGTACAGGCGCGGTGGAAGAGACCCC 1186
OY 961 CAGAGCAGCTGAGAGAGCAAGAGACCCATGCTGATTTCTTGAAGCGGAGTCCCTGAG 1020
DB 1187 CAGAGCAGCTGAGAGAGCAAGAGACCCATGCTGATTTCTTGAAGCGGAGTCCCTGAG 1246
OY 1021 GGGCGGGCCCTGAGTCAAGAGGAGGAGCCGAGACCTGAGAGGAGCTTGGGTGCGGGGG 1080
DB 1247 GGGCGGGCCCTGAGTCAAGAGGAGGAGCCGAGACCTGAGAGGAGCTTGGGTGCGGGGG 1306
OY 1081 CCTGTTTTGGAGAGGCATCAGCTCACCGCAGACACAGTGGGGTCTGCTGGAGAGAGC 1140
DB 1307 CCTGTTTTGGAGAGGCATCAGCTCACCGCAGACACAGTGGGGTCTGCTGGAGAGAGC 1366
OY 1141 CGGAGCAGCGAAGTGGAGCTTCTGATCTCGGCTGCGGAACTACAGTCCCGCAGACAC 1200
DB 1367 CGGAGCAGCGAAGTGGAGCTTCTGATCTCGGCTGCGGAACTACAGTCCCGCAGACAC 1426
OY 1201 TTCTACTGCTGGTGTCCAAAGGATGATATGTAG 1233
DB 1427 TTCTACTGCTGGTGTCCAAAGGATGATATGTAG 1459

RESULT 3
AX058610
LOCUS AX058610 1659 bp DNA linear PAT 17-JAN-2001
DEFINITION Sequence 17 from Patent W00075321.
ACCESSION AX058610
VERSION AX058610.1 GI:12310952
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1659)

LOCUS	DEFINITION	Mus musculus. Similar to single Ig IL-1R-related protein, clone
LOCUS BC010806	BC010806	1568 bp mRNA linear ROD 17-JUL-2001
61	AGGCTGCTGGGAGCTCAGTGGCTGTGAATCTACAGGCTGGGTAGTCTCTGGGCC	120
154	AGGCTGCTGGGAGCTCAGTGGCTGTGAATCTACAGGCTGGGTAGTCTCTGGGCC	223
121	CACCTGCTCCCTGCTTCAGTGCAGTGGCTGGAAGACGGGCTTCATTTAGGAATTGGGGC	180
224	CACCTGCTCCCTGCTTCAGTGCAGTGGCTGGAAGACGGGCTTCATTTAGGAATTGGGGC	283
181	CACCTACACCTTCACAGTACTCTGGGTCAGAGGCCAACCTGTTCAGAGCTGTGTTC	240
284	CACCTACACCTTCACAGTACTCTGGGTCAGAGGCCAACCTGTTCAGAGCTGTGTTC	343
241	AGTGCCTGGGGGTCAGAGTGCAGACCTGAAGCTTTAGGGGCTTCACCTGCTGCATTC	300
344	AGTGCCTGGGGGTCAGAGTGCAGACCTGAAGCTTTAGGGGCTTCACCTGCTGCATTC	403
301	CAGAACATCAGCTTCTCCTCTTCACTCTTACAGAGCTGGCCCTACAGACACGTGGCT	360
404	CAGAACATCAGCTTCTCCTCTTCACTCTTACAGAGCTGGCCCTACAGACACGTGGCT	463
361	GGGCTGCTGGCCCTCCCTGCTGCTGCTGGCCCTGCTGCTGGCCCTGCTGATTC	420
464	GGGCTGCTGGCCCTCCCTGCTGCTGCTGGCCCTGCTGCTGGCCCTGCTGATTC	523
421	AAGTGCCTTCACAGTACTCTGTGTACAGAGACGGGTATGGGAGGTGAGATTAAC	480
524	AAGTGCCTTCACAGTACTCTGTGTACAGAGACGGGTATGGGAGGTGAGATTAAC	583
481	GACGGAGAGCTCTACAGAGCTACGCTCTACAGAGCTCCCGAGAGACCGCAATTC	540
584	GACGGAGAGCTCTACAGAGCTACGCTCTACAGAGCTCCCGAGAGACCGCAATTC	643
541	GTGAATCTTCATCTAAAGCCGACCTGAGAGCGGCTGCGGCTACAACTCTTCTTGAC	600
644	GTGAATCTTCATCTAAAGCCGACCTGAGAGCGGCTGCGGCTACAACTCTTCTTGAC	703
601	GACCGGACCTCTGCGCGGCTGAGAGCCCTTCGCGCACCTTGTGTAACTGAGCCGC	660
704	GACCGGACCTCTGCGCGGCTGAGAGCCCTTCGCGCACCTTGTGTAACTGAGCCGC	763
661	TGCGGACCTCTCATGCTGTGTGCGAGCGCTTCAGAGCGGGGCTGGGTGACGCAC	720
764	TGCGGACCTCTCATGCTGTGTGCGAGCGCTTCAGAGCGGGGCTGGGTGACGCAC	823
721	AGCTTCCGGGAGGGCTGTGCGGCTGTGAGAGCTCACCGCAGACCCATCTTCAACC	780
824	AGCTTCCGGGAGGGCTGTGCGGCTGTGAGAGCTCACCGCAGACCCATCTTCAACC	883
781	TTTCAGAGGCTCAGAGCGGACCCCGCGCATCCCGCTGCTGCTGCCAGCACCGC	840
884	TTTCAGAGGCTCAGAGCGGACCCCGCGCATCCCGCTGCTGCTGCCAGCACCGC	943
841	CACCTGAGACCTCTCTCTGAGAGCGGCTCCGAGACCTTCTCCGATTTTGG	900
944	CACCTGAGACCTCTCTCTGAGAGCGGCTCCGAGACCTTCTCCGATTTTGG	1003
901	AAAGAGTGCAGCTGGCCTCCCGGAGAGTGCGGTACAGGCCGTTGGAAGAGAACCC	960
1004	AAAGAGTGCAGCTGGCCTCCCGGAGAGTGCGGTACAGGCCGTTGGAAGAGAACCC	1063
961	CAGAGCAGCTGCAGAGCAGACAAGGACCCCATGCTGATTTTCAGAGCCGAGTCCCTGAG	1020
1064	CAGAGCAGCTGCAGAGCAGACAAGGACCCCATGCTGATTTTCAGAGCCGAGTCCCTGAG	1123
1021	GGCGGGGCTGAGATCAGAGGTGAGACCGGACCTTGAGGGGAGACTGGGTGTC	1075
1124	GGCGGGGCTGAGATCAGAGGTGAGACCGGACCTTGAGGGGAGACTGGGTGTC	1178

ACCESSION	MGC:18899 IMAGE:4240425, mRNA, complete cds.
VERSION	BC010806
KEYWORDS	BC010806.1 GI:14789848
SOURCE	MGC.
ORGANISM	house mouse. Mus musculus
REFERENCE	Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1568)
AUTHORS	Strausberg, R.
TITLE	Direct Submission
JOURNAL	Submitted (12-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk Email: gcgaps@mail.nih.gov Tissue Procurement: Jeffrey E. Green, M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: http://www-snpcc.stanford.edu Contact: (Dickson, Mark) mcd@pacil.stanford.edu Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAK Plate: 25 Row: P Column: 10 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.
SOURCE	Location/Qualifiers 1..1568 /organism="Mus musculus" /db_xref="taxon:10090" /clone="MGC:18899 IMAGE:4240425" /tissue_type="Kidney, normal, 5 month old male mouse." /clone_id="NCI CGAP_Kid14" /lab_host="DH10B" /note="Vector: pCMV-SPORT6" length=1349 /codon_start=1 /product="Similar to single Ig IL-1R-related protein" /protein_id="AAH10806.1" /db_xref="GI:14789849" /translation="MAGVCDAWPNFLSPSDQALGLAREVALNCTAWFSRPQCPC PSYOMLKDGLANGSHFSLHEDFWANSSEIVSVLVNLTMAEDYGFETCSVMNVN SSHFETLRAGPAGHVAVALSLVLVLTALLVAYLCRNMLMTADTYGEVENND GKLYDAVYSDDCEDERKFNFILPKLETRRGKKLPLEDIDLPRAPSADLVNLS RCRLATVLDAPLSRMCGSCFPREGCRILETFRRPIFTFEQQRREPITPARLLRL OHRRLVLVLMKPESVTSPSSDFKELOLAPRKVOYRPVEDPOTIRQQDDPEVLVR GRQAQGGMSELDPDEGDIGRVGPVEPPPLDETFRICIGESHSEMVDYSLGSRS NYSKRDFCYLSEDDV"
CDS	BASE COUNT 359 a 433 c 345 t ORIGIN
Query Match	69.8% Score 860.2; DB 10; Length 1568;
Best Local Similarity	81.7%; Pred. No. 6.5e-122;
Matches 1007; Conservative	0; Mismatches 223; Indels 3; Gaps 1;
OY	1 ATGGCAGGTGTCGTGAATGAGGCCCCCTGACTTTCCCTCCCGCTGGAAGACGAGTGCTG 60
Db	120 ATGGCAGGTGTCGTGAATGAGGCCCCCTTAATTCTTTCCCATCTGGAAGACCAGGCTTG 179
OY	61 AGGCGTGTGGGGAGCTCACTGAGCTTGAATGACAGCGCTTGAGTCTTG6GGCCC 120
Db	180 GGCTTGCCCTTGGCAGAGAAGATTGCTTGAATTGACACAGCTTGGGTGTTCTTAGGGCCC 239
OY	121 CACTGCTCTGCTTCACTGATCCAGTGGCTGAAGAAGCGGCTTCCATTGGAAATTGGGGCC 180

Db 240 CAGTGTCCCGCAGCATCAGTACGAGTGGCTGAAGAATGAGTGGCATTTGGGCATGGAAGC 299
 Oy 181 CACTACAGCCTCCAGAGTACCTGCGGTCAGAGCCAACTGTCAGAGTGTGTGTC 240
 Db 300 CACTACAGCCTCCAGAGTACCTGCGGTCAGAGCCAACTGTCAGAGTGTGTGTC 356
 Oy 241 AGTGTCTGGGGTCAAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 300
 Db 357 AGTGTCTGGGGTCAAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 416
 Oy 301 CAGAACATGCTTCT 360
 Db 417 TGAAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 476
 Oy 361 GCGGTGTGCT 420
 Db 477 GCAATACAGTCT 536
 Oy 421 AAGTGTCT 480
 Db 537 AAGTGTCT 596
 Oy 481 GAGGGAAGTCTACAGAGCCTACAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
 Db 597 GATGGAGATTTATACAGTCT 656
 Oy 541 GTGAACTTCT 600
 Db 657 GTAAATTTCT 716
 Oy 601 GACCGCAGCT 660
 Db 717 GACCGCAGCT 776
 Oy 661 TGCGAGCCT 720
 Db 777 TGCGAGCCT 836
 Oy 721 AGTGTCTGGGGTCAAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 780
 Db 837 AGTGTCTGGGGTCAAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 896
 Oy 781 TTGAGGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
 Db 897 TTGAGGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 956
 Oy 841 CACCTGTGACCTTCT 900
 Db 957 CACCTGTGACCTTCT 1016
 Oy 901 AAGAAGTGTGACCTTCT 960
 Db 1017 AAGAAGTGTGACCTTCT 1076
 Oy 961 CAAAGAGTGTGACCTTCT 1020
 Db 1077 CAAAGAGTGTGACCTTCT 1136
 Oy 1021 GGGCGGGGCT 1080
 Db 1137 GGGCGGGGCT 1196
 Oy 1081 CTTGTCTTGTGAGAGCCTTCT 1140
 Db 1197 CTTGTCTTGTGAGAGCCTTCT 1256
 Oy 1141 CAGAGCAGCAGAGTGTGAGAGCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1200
 Db 1257 CAGAGCAGCAGAGTGTGAGAGCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1316
 Oy 1201 TTCTACTGCTGTGTCAGAGTGTGTCAGAGTGTGTCAGAGTGTGTCAGAGTGTGTC 1233
 Db 1317 TTCTACTGCTGTGTCAGAGTGTGTCAGAGTGTGTCAGAGTGTGTCAGAGTGTGTC 1349

RESULT 6
 AF239957
 LOCUS AF239957 1421 bp mRNA linear ROD 21-MAR-2000
 DEFINITION Mus musculus single Ig IL-1R-related protein mRNA, complete cds.
 ACCESSION AF239957
 VERSION AF239957.1 GI:7271942
 KEYWORDS house mouse, Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1421)
 Poltorak, A. and Beutler, B.
 Molecular cloning of mouse SIGIRR, an orphan receptor with a Toll-like cytoplasmic domain
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1421)
 Poltorak, A. and Beutler, B.
 Direct Submission
 TITLE Submitted (01-MAR-2000) Internal Medicine, UT Southwestern Medical Center/HMUT, 5323 Harry Hines Boulevard, Dallas, TX 75335-9050, USA
 FEATURES
 source
 1. 1421
 /organism="Mus musculus"
 /strain="BALB/c"
 /db_xref="taxon:10090"
 /cell_line="RAW 264.7"
 109..1338
 /note="SIGIRR: Toll-like cytoplasmic domain"
 /codon_start=1
 /product="single Ig IL-1R-related protein"
 /protein_id="AA04701.1"
 /db_xref="GI:7271943"
 /translating="MAGCDMAAPNLSSEDOALGALREVALNCTANVAFSPQCPRO
 PSYOMKDLALGNGSHSLHEDVPSANSEIYSVLYNLTNADYCPNCSYANV
 SSSPTIRASGAGHAYAAVLASLVLYLVLLVALLVYKCLNMLNLYODTYGEVEND
 GKIKDVAIVYSQCPEDRKRVNLLKQVLEDRLEKRLRILFTFGQREPLHPLNLLR
 KCRLLVLYLSDFLSRPMQSFQREGLDLELRLRILFTFGQREPLHPLNLLR
 OHRLVTLVLMKRFQSVTPSSDWEKQLOALPRKQVHAPVGDQFTRLDQDKMLTVR
 GRAQGRGMSLELDPDGLDVRKPEVPEPPLVQETRICDSHSGSEMDVLSGR
 NYSARDYCLVSDV"
 BASE COUNT 280 a 406 g 324 t 4 others
 ORIGIN
 Query Match 69.5%; Score 856.6; DB 10; Length 1421;
 Best Local Similarity 81.5%; Pred. No. 2,4e-121;
 Matches 1004; Conservative 0; Mismatches 225; Indels 3; Gaps 1;

[illegible]

REFERENCE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	Penton-RoJ.G., Polentarutti,N., Muzio,M. and Mantovani,A.
TITLE	TIR8: a novel toll/interleukin-1 receptor family member
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 1540)
AUTHORS	Penton-RoJ.G., Polentarutti,N., Muzio,M. and Mantovani,A.
TITLE	Direct Submission
JOURNAL	Submitted (14-DEC-1998) Immunology, Mario Negri Institute, Via Erlrate 62, Milano, MI I-20152, Italy
FEATURES	
source	Location/Qualifiers
	1..1540
gene	/organism="Mus musculus"
	/db_xref="taxon:10090"
	1..1540
CDS	/gene="TIR8"
	115..1344
	/gene="TIR8"
	/note="putative transmembrane receptor of the Toll/IL-1 family; cytoplasmic domain similar to other IL-1R family members"
	/codon_start=1
	/product="toll/interleukin-1 receptor 8"
	/protein_id="AAF26200.1"
	/db_xref="GI:6714888"
	/translation="MAGCYDAPMFSLPSDEDALGLAREVALNCTAMVFSROCPQ PSYVMKDGALGNGSHFSLHEDFWANSEIVSYVLNLNAEDYGFTSVWV GSKETLRMRGPAGHVAVALSLVLVLALLVKRLMLMWOTDYREVENND GSLFDATVYSDDCEPDRFAVEILKPOLERRRGKFLIEDRLFRADASADLVNS RCRLIIVLSDAFLSPMCQSFEDEDAAYWSSPADLSSTRARGVPTLLSGSR HDRRLITVLMKPGSVSPSSDFKELOLAIPRKQVYPVGSDPTRLODDKPLLYR GRAOGKGMESELDPDEBDLGAVGFGEPIPLQLERLICIGESHSEMDVSUGSR NYSRTPFYCLVSSDDV"
BASE COUNT	334 a 434 c 429 g 343 t
ORIGIN	
Query Match	67.0%; Score 826; DB 10; Length 1540;
Best Local Similarity	81.0%; Pred. No. 1,1e-116;
Matches 999: Conservative	0; Mismatched 230; Indels 5; Gaps 3;
Oy	1 ATGCCAGGTGTCGTGAAGGGCCCTCAGTCCTCTCCCGTCGAAGACAGTGCTG 60
Dd	115 ATGGCAGGTGTCGTGAATGACCCTTAATTCTTTCCCATCTGAAGACAGGCTTG 174
Oy	61 AGGCTGCCTTGGGACGCTCAGTGCTGCAATGCAAGGCTTGAGTAGCTCGGGCCC 120
Dd	175 GGTCTGGCCCTTGGGAGAAGAGTCTTTGAATTGCAAGCTTGGCTTCTGTACAGCCC 234
Oy	121 CACTGCTCCCTGCTTCACGTCCAGTGCTGAAGAAGCGGGCTTCATTGGAAATTGGGGGC 180
Dd	235 CAGTGTCCCCAGCCATCAGTGCAGTGCGCTGAAGAAGTGGCTGGCGCAATGGAAGGC 294
Oy	181 CACTCAGCCTCCACAGATACCTCCGGGTCAAAGGCCAACCTGTGACAGAGCTCTGTGCGC 240
Dd	295 CACTCAGCCTCCACAGAGACTCTCGGGTCAAGCGCAACTTCTCAGA--GATTGTGTC 351
Oy	241 AGTGTCTGGGGGTCAAGCTGACAGCACTGAAGTCTATGGGGGCTTCACCTGCTCATC 300
Dd	352 AGTGTCTGGGGGTCAACTGTGACCAATGCAAGAGACTGTGAACCTTACCTGTTCTGTC 411
Oy	301 CAGAATCATGCTTCTCTCTCTTCACTCTTCAAGAGAGCTGGGCCCTTACAAGCCACGTGGCT 360
Dd	412 TGGATGTGATGCCCATCTTCTTCACTCTTTGGGAGCGCCGCTGTGGGCAATGTGCT 471
Oy	361 GGCGTGTGGGCTCTCCCTCTGCTGCTGCTGGCCCTGCTGCTGCTGCTCATATGC 420
Dd	472 GCAGTACGTGGCTCTCCCTCTGGTCTGTGTTGTGTGTGCTGCTGCTCTCATATT 531
Oy	421 AAGTCCCTTCACAGCTGCTCTGTGATCACGAGAGCGGTATGGGAGGTGGAGATAAAC 480
Dd	532 AAGTGTGGGTGACATCTCTCTTTGGATCAACAGACATTACGGGAGGTGGAGATGAAC 591
Oy	481 GACGGGAAGCTTCAGCAGCCTTACGCTTCMACAGCAGACTGCCCGAGAGACCGCAAGTTC 540

Db 592 GATGGAGAGTATACGATGCTAGCTGCTCTATAGAGTCCGACAGAGACCGAATTT 651
 QY 541 GAGACCTTCATCTAAAGCCGACGCTGAGAGAGGCGGCGTACAGACCTTCCTGAC 600
 Db 652 GTAATTTTATTCGAAAGCTTCAGATGAGAGGCGGCGTACAGACCTTCCTGAC 711
 QY 601 GACCCGACCTTCGCGCGGCTGAGACCTTCGCGGACCTTCGAGACCTGACCGC 660
 Db 712 GACCCGACCTTCGCGCGGCTGAGACCTTCGCGGACCTTCGAGACCTGACCGC 771
 QY 661 TGCCGACGCTTCATGCTGCTTCGAGAGGCGGCGTTCGAGAGGCGGCGTTCGAGAG 720
 Db 772 TGCTGCGCTTCATGCTGCTTCGAGAGGCGGCGTTCGAGAGGCGGCGTTCGAGAG 831
 QY 721 AGCTTCGCGGACGCTGCGCGGCTGAGAGGCGGCGTTCGAGAGGCGGCGTTCGAGAG 780
 Db 832 AGCTTCGCGGACGCTGCGCGGCTGAGAGGCGGCGTTCGAGAGGCGGCGTTCGAGAG 890
 QY 781 TTGAGAGGCGGACGCGGCGGCGTTCGAGAGGCGGCGTTCGAGAGGCGGCGTTCGAGAG 839
 Db 891 TTGAGAGGCGGACGCGGCGGCGTTCGAGAGGCGGCGTTCGAGAGGCGGCGTTCGAGAG 950
 QY 840 CCACCTGCTGACCTTCTGCTGCTGAGAGGCGGCGTTCGAGAGGCGGCGTTCGAGAG 899
 Db 951 CCACCTGCTGACCTTCTGCTGCTGAGAGGCGGCGTTCGAGAGGCGGCGTTCGAGAG 1010
 QY 900 GAAAGAGTGCAGCTGCGCGGCGGCGTTCGAGAGGCGGCGTTCGAGAGGCGGCGTTCGAGAG 959
 Db 1011 GAAAGAGTGCAGCTGCGCGGCGGCGTTCGAGAGGCGGCGTTCGAGAGGCGGCGTTCGAGAG 1070
 QY 960 CCAGAGGAGCTGAGAGGCGGCGGCGTTCGAGAGGCGGCGTTCGAGAGGCGGCGTTCGAGAG 1019
 Db 1071 TCAGAGGAGCTGAGAGGCGGCGGCGTTCGAGAGGCGGCGTTCGAGAGGCGGCGTTCGAGAG 1130
 QY 1020 GGGCGGCGGCGTTCGAGAGGCGGCGGCGTTCGAGAGGCGGCGTTCGAGAGGCGGCGTTCGAGAG 1079
 Db 1131 GGGCGGCGGCGTTCGAGAGGCGGCGGCGTTCGAGAGGCGGCGTTCGAGAGGCGGCGTTCGAGAG 1190
 QY 1080 GCGCTTTTGGAGAGGCGGCGGCGTTCGAGAGGCGGCGTTCGAGAGGCGGCGTTCGAGAG 1139
 Db 1191 ACCTTCTTTTGGAGAGGCGGCGGCGTTCGAGAGGCGGCGTTCGAGAGGCGGCGTTCGAGAG 1250
 QY 1140 CCGAGAGGCGGCGGCGGCGGCGTTCGAGAGGCGGCGTTCGAGAGGCGGCGTTCGAGAG 1199
 Db 1251 CCAGAGGAGCTGAGAGGCGGCGGCGTTCGAGAGGCGGCGTTCGAGAGGCGGCGTTCGAGAG 1310
 QY 1200 CTTTACTGCTGCTGCTGCTGAGAGGCGGCGTTCGAGAGGCGGCGTTCGAGAGGCGGCGTTCGAGAG 1233
 Db 1311 CTTTACTGCTGCTGCTGCTGAGAGGCGGCGTTCGAGAGGCGGCGTTCGAGAGGCGGCGTTCGAGAG 1344
 REPEAT 8
 AC108908
 LOCUS AC108908 61326 bp DNA linear HTG 01-FEB-2002
 DEFINITION Mus musculus clone RP23-179K7, LOW-PASS SEQUENCE SAMPLING.
 AC108908
 VERSION AC108908.1 GI:18464144
 KEYWORDS HTG; HTGS_PHASE0.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 REFERENCE
 AUTHORS 1 (bases 1 to 61326) Nusbaum, C. and Lander, E.
 TITLE Mus musculus, clone RP23-179K7
 JOURNAL Unpublished
 AUTHORS 2 (bases 1 to 61326) Birren, B., Linton, L., Lander, E., Allen, N.,
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N.,
 Anderson, S., Barta, N., Bastien, V., Boguslavsky, L., Bouhassira, B.,
 Brown, A., Camarata, J., Campio, A., Chang, J., Charaz, B.,
 Cooke, P., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
 Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,

TITLE
JOURNAL
COMMENT

Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gird, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hages, B., Horton, L., Hulme, W., Iliev, K., Johnson, R., Jones, C.,
 Kamal, A., Karatas, A., Kellis, C., Lachowicz, K., Lamata, R.,
 Landers, J., Leniczky, J., Levine, R., Liu, G., Maclean, C.,
 MacDonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M.,
 McEwan, P., McKernan, K., Melnick, J., Meneus, L., Mihova, T.,
 Mlenka, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
 Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
 Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,
 Retta, R., Riaback, M., Riley, R., Rhee, C., Rogov, P., Roman, J.,
 Roselli, M., Roy, A., Santos, R., Schuer, S., Schuback, R., Seaman, S.,
 Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talmas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Tridillo, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
 Zainou, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (01-FEB-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L19209
 Center clone name: 179_K7

 * NOTE: This record contains 78 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.
 * 1 678: contig of 678 bp in length
 * 679 778: gap of 100 bp
 * 779 1415: contig of 637 bp in length
 * 1416 1515: gap of 100 bp
 * 1516 2195: contig of 680 bp in length
 * 2196 2295: gap of 100 bp
 * 2296 2983: contig of 688 bp in length
 * 2984 3083: gap of 100 bp
 * 3084 3769: contig of 686 bp in length
 * 3770 3869: gap of 100 bp
 * 3870 4573: contig of 704 bp in length
 * 4574 4673: gap of 100 bp
 * 4674 5367: contig of 694 bp in length
 * 5368 5467: gap of 100 bp
 * 5468 6161: contig of 694 bp in length
 * 6162 6261: gap of 100 bp
 * 6262 6947: contig of 686 bp in length
 * 6948 7047: gap of 100 bp
 * 7048 7727: contig of 680 bp in length
 * 7728 7827: gap of 100 bp
 * 7828 8512: contig of 685 bp in length
 * 8513 8612: gap of 100 bp
 * 8613 9297: contig of 685 bp in length
 * 9298 9397: gap of 100 bp
 * 9398 10094: contig of 697 bp in length
 * 10095 10194: gap of 100 bp
 * 10195 10892: contig of 698 bp in length
 * 10893 10992: gap of 100 bp
 * 10993 11674: contig of 682 bp in length
 * 11675 11774: gap of 100 bp

11775 12457: contig of 683 bp in length
12458 12557: gap of 100 bp
13258 13238: contig of 681 bp in length
13239 13338: gap of 100 bp
13339 14002: contig of 664 bp in length
14003 14102: gap of 100 bp
14103 14787: contig of 685 bp in length
14788 14887: gap of 100 bp
14888 15574: contig of 687 bp in length
15575 15674: gap of 100 bp
15675 16357: contig of 683 bp in length
16358 16457: gap of 100 bp
16458 17134: contig of 677 bp in length
17135 17234: gap of 100 bp
17235 17920: contig of 686 bp in length
17921 18020: gap of 100 bp
18021 18706: contig of 686 bp in length
18707 18806: gap of 100 bp
18807 19501: contig of 695 bp in length
19502 19601: gap of 100 bp
19602 20302: contig of 701 bp in length
20303 20402: gap of 100 bp
20403 21095: contig of 693 bp in length
21096 21195: gap of 100 bp
21196 21880: contig of 685 bp in length
21881 21980: gap of 100 bp
21981 22670: contig of 690 bp in length
22671 22770: gap of 100 bp
22771 23445: contig of 675 bp in length
23446 23545: gap of 100 bp
23546 24232: contig of 687 bp in length
24233 24332: gap of 100 bp
24333 25008: contig of 676 bp in length
25009 25108: gap of 100 bp
25109 25800: contig of 692 bp in length
25801 25900: gap of 100 bp
25901 26602: contig of 702 bp in length
26603 26702: gap of 100 bp
26703 27385: contig of 683 bp in length
27386 27485: gap of 100 bp
27486 28169: contig of 684 bp in length
28170 28269: gap of 100 bp
28270 28944: contig of 675 bp in length
28945 29044: gap of 100 bp
29045 29735: contig of 691 bp in length
29736 29835: gap of 100 bp
29836 30515: contig of 680 bp in length
30516 30615: gap of 100 bp
30616 31289: contig of 674 bp in length
31290 31389: gap of 100 bp
31390 32069: contig of 680 bp in length
32070 32169: gap of 100 bp
32170 32856: contig of 687 bp in length
32857 32956: gap of 100 bp
32957 33653: contig of 697 bp in length
33654 33753: gap of 100 bp
33754 34452: contig of 699 bp in length
34453 34552: gap of 100 bp
34553 35256: contig of 704 bp in length
35257 35356: gap of 100 bp
35357 36056: contig of 700 bp in length
36057 36156: gap of 100 bp
36157 36851: contig of 695 bp in length
36852 36951: gap of 100 bp
36952 37633: contig of 682 bp in length
37634 37733: gap of 100 bp
37734 38427: contig of 694 bp in length
38428 38527: gap of 100 bp
38528 39201: contig of 674 bp in length
39202 39301: gap of 100 bp
39302 39985: contig of 684 bp in length
39986 40085: gap of 100 bp
40086 40755: contig of 670 bp in length

40756 40855: gap of 100 bp
40856 41535: contig of 680 bp in length
41536 41635: gap of 100 bp
41636 42339: contig of 704 bp in length
42340 42439: gap of 100 bp
42440 43123: contig of 684 bp in length
43124 43223: gap of 100 bp
43224 43905: contig of 682 bp in length
43906 44005: gap of 100 bp
44006 44686: contig of 681 bp in length
44687 44786: gap of 100 bp
44787 45483: contig of 697 bp in length
45484 45583: gap of 100 bp
45584 46266: contig of 683 bp in length
46267 46366: gap of 100 bp
46367 47049: contig of 683 bp in length
47050 47149: gap of 100 bp
47150 47834: contig of 685 bp in length
47835 47934: gap of 100 bp
47935 48623: contig of 689 bp in length
48624 48723: gap of 100 bp
48724 49408: contig of 685 bp in length
49409 49508: gap of 100 bp
49509 50213: contig of 705 bp in length
50214 50313: gap of 100 bp
50314 51024: contig of 711 bp in length
51025 51124: gap of 100 bp
51125 51800: contig of 676 bp in length
51801 51900: gap of 100 bp
51901 52591: contig of 691 bp in length
52592 52691: gap of 100 bp
52692 53436: contig of 745 bp in length
53437 53536: gap of 100 bp
53537 54215: contig of 679 bp in length
54216 54315: gap of 100 bp

Query Match 11.5% Score 141.6; DB 2; Length 61326;
Best Local Similarity 71.4%; Pred. 7.8e-13;
Matches 200; Conservative 0; Mismatches 79; Indels 1; Gaps 1;

QY 857 TCGCTGAGAGCCCGGCTCGTGACTCTTCCTCCGATTTTGGAAAGATGCACTGG 916
DB 30061 TTCTGTCCAAACCTCGGCCCCACAGACTCTTCCTGATTTTGGAAAGATGCACTAG 30120
QY 917 CGGTGCGGGAAGTGGCGGTACAGCCGCTGGAAGACACCCCGACGACCTGAG 976
DB 30121 CACTGCGACGGAAGTGCAGTACAGCCGCTGGAAGACACCCCTCAACCCGACTTCAGG 30180
QY 977 ACGACAGAGACCCCATGATTCCTTCGAGGCGAGTCCCTGAGGCGCGGCGCTGAGCT 1036
DB 30181 ATGACAAAGATCCCATGCTAATGCTGAGAGAGCTCTCCCGGCGGCGATGAGCT 30240
QY 1037 CAGAGGTGACCCCGACCTGAGGCGAGCTGCGGCTCGCGGGGCTGTTTGGAGAC 1096
DB 30241 CAGAGCTGATCCAGACCTGAGGCGAGACCTGATGCTACATGCAAGTCAAGGGTCA 30300
QY 1097 CA-TGAGCTCCACCGACACACACTGGGCTCGCTGGAG 1135
DB 30301 AACACAGCTCCAGCTGCTGGAAGAGCTCAGGCTTGGAG 30340

RESULT 9 G38549 324 bp DNA linear SRS 23-JUN-1998
G38549/c SHG-64262 Human Homo sapiens STS genomic, sequence tagged site.
DEFINITION G38549
ACCESSION G38549.1 GI:3249323
VERSION G38549.1
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 324)

us-09-598-443-1.rge

Contact: Richard M. Weers
Stanford Human Genome Center (SHGC)
Stanford University School of Medicine
Department of Genetics, M-344, Stanford, CA 94305, USA
Tel.: 4157259687
Fax: 4157259689
Email: richard.weers@stanford.edu
URL: <http://www.genetics.stanford.edu>
Primer A: AGGAGGAGGCTGCTGGTAGCA
Primer B: ACTGCTGTGTCTCCAGAGAT
STS size: 111
PCR Profile:

Denaturation:	95 degrees C for 10 minutes
Annealing:	64 degrees C for 30 seconds
Polymerisation:	72 degrees C for 30 seconds
PCR Cycles:	72 degrees C for 23 seconds
Thermal Cycler:	30
Protocol:	Pertkin Elmer 9700
Template:	
Primer:	25 ng
divers:	each 1 uM
Amplifying Gold Polymerase:	each 200 uM
Total Vol:	0.07 units/uM 5 uL

EST sequences developed for scoring on agarose gels. Designed and developed at the Stanford Human Genome Center on Cancer Genome Anatomy Project sequences. sequence AA490620.
1..324

		65	t	22	others
ORGANISM="Homo sapiens"					
/db_xref="taxon:9606"					
map="11"					
11one_11b-"Human"					
121 . .231					
complement(1212 . .231)					
93 c	50	a	94	g	65
BASE COUNT					
primer_bind					
primer_bind					
STS					
ORIGIN					

	Match	9.9%	Score 122	DB 11	Length 324
Local Similarity	100.0%				
Matches 122	Conservative	0	Mismatches 0	Indels 0	Gaps 0
OY	1112 ACACCAATGGGGTCTCCCTGGAGAGACCGCGGACGACGAATGAGATCTCGAATCTG				1171
Db	324 ACACGAGGGGGTCTGCTGGAGAGACCGCGGACGACGAATGAGATCTCGAATCTG				
OY	1172 GCTCGCAAACTCAAGTGCCTCCACAGACTTCTACTGCTCTGGTCCCAAGATGATATG				265
Db	264 GCTCGCAAACTCAAGTGCCTCCACAGACTTCTACTGCTCTGGTCCCAAGATGATATG				1231
OY	1232 AG 1233				
Db	204 AG 203				

RESULT 10	
AC109272	
LOCUS	
DEFINITION	AC109272 Mus musculus clone RP23-354L18, DNA linear
ACCESSION	AC109272
VERSION	AC109272.1 GI:18482193
	HTG 03-FEB-2002
	SEQUENCE SAMPLING.

KEYWORDS	HTG; HTGS. PHASED.
SOURCE	house mouse.
ORGANISM	Mus musculus
	Musculi
	Kukariyola, Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Rodentia; Sciurognathi: Muridae: Murinae: Mus (bases 1 to 62831)
REFERENCE AUTHORS	Birren, B., Lincoln, L., Nussbaum, C. and Lander, E.
TITLE	Unpublished
JOURNAL	Unpublished, clone RP23-354118
REFERENCE	2 (bases 1 to 62831)
AUTHORS	Birren, B., Lincoln, L., Nussbaum, C. and Lander, E.

TITLE
 JOURNAL
 Direct Submission
 Submitted (03-FEB-2002) Whitehead Institute/MIT Center for Genome
 Research, 520 Charles Street, Cambridge, MA 02141, USA
 COMMENT
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
 ----- Genome (Genome)

```

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WMR
Web site: http://www.seq.wi.mit.edu
Contact: sentence.submissions@genome.wi.mit.edu
Project Information
Center project name: L20944
Center clone name: 354_L_18

NOTE: This record contains 79 individual
sequencing reads that have not been assembled into
contigs. Runs of N are used to separate the reads
and the order in which they appear is completely
arbitrary. Low-pass sequence sampling is useful for
identifying low-pass relationships that may be gene-rich and allows
overlapping relationships among clones to be deduced.
However, it should not be assumed that this clone
will be sequenced to completion. In the event that
the record is updated, the accession number will
be preserved.

1
687: contig of 687 bp in length
788: gap of 100 bp
1505: contig of 717 bp in length
1505: 1604: gap of 100 bp
1605: 2300: contig of 696 bp in length
2301: 2400: gap of 100 bp
2401: 3116: contig of 716 bp in length
3117: 3216: gap of 100 bp
3217: 3865: contig of 649 bp in length
3866: 3965: gap of 100 bp
3966: 4680: contig of 715 bp in length
4681: 4780: gap of 100 bp
4781: 5484: contig of 704 bp in length
5485: 5584: gap of 100 bp
6288: contig of 704 bp in length

```

```
* 6289 6388: gap of 100 bp
* 6389 7090: contig of 702 bp in length
* 7091 7190: gap of 100 bp
* 7191 7906: contig of 716 bp in length
* 7907 8006: gap of 100 bp
* 8007 8709: contig of 703 bp in length
* 8710 8809: gap of 100 bp
* 8810 9507: contig of 698 bp in length
* 9508 9607: gap of 100 bp
* 9608 10316: contig of 709 bp in length
* 10317 10416: gap of 100 bp
* 10417 11121: contig of 705 bp in length
* 11122 11221: gap of 100 bp
* 11222 11926: contig of 705 bp in length
* 11927 12026: gap of 100 bp
* 12027 12735: contig of 709 bp in length
* 12736 12835: gap of 100 bp
* 12836 13525: contig of 690 bp in length
* 13526 13625: gap of 100 bp
* 13626 14329: contig of 704 bp in length
* 14330 14429: gap of 100 bp
* 14430 15135: contig of 706 bp in length
* 15136 15235: gap of 100 bp
* 15236 15943: contig of 708 bp in length
* 15944 16043: gap of 100 bp
* 16044 16746: contig of 703 bp in length
* 16747 16846: gap of 100 bp
* 16847 17524: contig of 678 bp in length
* 17525 17624: gap of 100 bp
* 17625 18325: contig of 701 bp in length
* 18326 18425: gap of 100 bp
* 18426 19136: contig of 711 bp in length
* 19137 19236: gap of 100 bp
* 19237 19930: contig of 694 bp in length
* 19931 20030: gap of 100 bp
* 20031 20733: contig of 703 bp in length
* 20734 20833: gap of 100 bp
* 20834 21538: contig of 705 bp in length
* 21539 21638: gap of 100 bp
* 21639 22342: contig of 704 bp in length
* 22343 22442: gap of 100 bp
* 22443 23152: contig of 710 bp in length
* 23153 23252: gap of 100 bp
* 23253 23940: contig of 688 bp in length
* 23941 24040: gap of 100 bp
* 24041 24742: contig of 702 bp in length
* 24743 24842: gap of 100 bp
* 24843 25539: contig of 697 bp in length
* 25540 25639: gap of 100 bp
* 25640 26338: contig of 699 bp in length
* 26339 26438: gap of 100 bp
* 26439 27137: contig of 699 bp in length
* 27138 27237: gap of 100 bp
* 27238 27901: contig of 664 bp in length
* 27902 28001: gap of 100 bp
* 28002 28700: contig of 699 bp in length
* 28701 28800: gap of 100 bp
* 28801 29503: contig of 703 bp in length
* 29504 29603: gap of 100 bp
* 29604 30301: contig of 698 bp in length
* 30302 30401: gap of 100 bp
* 30402 31092: contig of 691 bp in length
* 31093 31192: gap of 100 bp
* 31193 31906: contig of 714 bp in length
* 31907 32006: gap of 100 bp
* 32007 32707: contig of 701 bp in length
* 32708 32807: gap of 100 bp
* 32808 33492: contig of 685 bp in length
* 33493 33592: gap of 100 bp
* 33593 34296: contig of 704 bp in length
* 34297 34396: gap of 100 bp
* 34397 35101: contig of 705 bp in length
* 35102 35201: gap of 100 bp
```

```
* 35202 35911: contig of 710 bp in length
* 35912 36011: gap of 100 bp
* 36012 36718: contig of 707 bp in length
* 36719 36818: gap of 100 bp
* 36819 37522: contig of 704 bp in length
* 37523 37622: gap of 100 bp
* 37623 38339: contig of 717 bp in length
* 38340 38439: gap of 100 bp
* 38440 39146: contig of 707 bp in length
* 39147 39246: gap of 100 bp
* 39247 39958: contig of 712 bp in length
* 39959 40058: gap of 100 bp
* 40059 40750: contig of 692 bp in length
* 40751 40850: gap of 100 bp
* 40851 41554: contig of 704 bp in length
* 41555 41654: gap of 100 bp
* 41655 42369: contig of 715 bp in length
* 42370 42469: gap of 100 bp
* 42470 43171: contig of 702 bp in length
* 43172 43271: gap of 100 bp
* 43272 43931: contig of 660 bp in length
* 43932 44031: gap of 100 bp
* 44032 44737: contig of 706 bp in length
* 44738 44837: gap of 100 bp
* 44838 45538: contig of 701 bp in length
* 45539 45638: gap of 100 bp
* 45639 46342: contig of 704 bp in length
* 46343 46442: gap of 100 bp
* 46443 47152: contig of 710 bp in length
* 47153 47252: gap of 100 bp
* 47253 47950: contig of 698 bp in length
* 47951 48050: gap of 100 bp
* 48051 48757: contig of 707 bp in length
* 48758 48857: gap of 100 bp
* 48858 49525: contig of 668 bp in length
* 49526 49625: gap of 100 bp
* 49626 50337: contig of 712 bp in length
* 50338 50437: gap of 100 bp
* 50438 51136: contig of 699 bp in length
* 51137 51236: gap of 100 bp
* 51237 51944: contig of 708 bp in length
* 51945 52044: gap of 100 bp
* 52045 52759: contig of 715 bp in length
* 52760 52859: gap of 100 bp
* 52860 53334: contig of 475 bp in length
* 53335 53434: gap of 100 bp
* 53435 54142: contig of 708 bp in length
* 54143 54242: gap of 100 bp
* 54243 54954: contig of 712 bp in length
* 54955 55054: gap of 100 bp
```

Query Match 9.2%; Score 113; DB 2; Length 62831;
Best Local Similarity 86.2%; Pred. No. 1.8e-08;
Matches 125; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

```
QY 724 TTCGGGAGGCGCTGCGGCTGCTGAGCTCAGCCGAGACCCATCTTCATCACTTC 783
DB 21277 TTCAGGAGGAGTGTGCGCTACTGAGCTCACCAGACTATCTTCACACCTTT 21336
QY 784 GAGGGCCAGAGGCGGACCCGCGGCGCTCCGCTGCTGGCCAGCACGCCAC 843
DB 21337 GAGGGCCAGAGGCGGTAGCCCATACACCTGCTCTCGGCTCTGCGCCAGCACGCCAC 21396
QY 844 CTGCTGACCTGCTGCTGAGGCG 868
DB 21397 CTCGTGACCCGTGCTTGGAGCC 21421
```

RESULT 11
AC108908/c AC108908 61326 bp DNA linear HTG 01-FEB-2002
LOCUS Mus musculus clone RP23-179K7, LOW-PASS SEQUENCE SAMPLING.
DEFINITION AC108908
ACCESSION

VERSION
 AC108908.1 GI:18464144
 HTG: HTGS PHASED.
 SOURCE
 house mouse
 ORGANISM
 Mus musculus
 REFERENCE
 Mammalia: Euteleostomi: Chordata: Craniata: Vertebrata: Euteleostomi:
 Mammalia: Euteleostomi: Rodentia: Sciurognathi: Muridae: Murinae: Mus.
 1 (bases 1 to 61326)
 AUTHORS
 Birren, B., Linton, L., Nusbaum, C. and Lander, E.
 TITLE
 Mus musculus, clone RP23-179K7
 JOURNAL
 Unpublished
 REFERENCE
 2 (bases 1 to 61326)
 AUTHORS
 Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,
 Brown, A., Camarata, J., Campione, V., Boguslavsky, L., Boukagalter, B.,
 Choe, P., Colangelo, M., Collins, S., Collamore, A., Cook, A.,
 Cooke, P., Dekrellano, K., Dewar, K., Diaz, J. S., Dodge, S., Fero, S.,
 Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Glade, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kanat, A., Karatas, A., Kells, C., Lacroque, K., Lamarca, R.,
 Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C.,
 MacDonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M.,
 McMan, P., McKenna, K., Melton, J., Meneus, L., Minova, T.,
 Mlenka, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
 Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
 Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,
 Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,
 Roselli, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S.,
 Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
 Tophan, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
 JOURNAL
 COMMENT
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
 Center Code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 Project Information
 Center Project name: U19209
 Center Clone name: 179_K7

* NOTE: This record contains 78 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1
 679 778: gap of 100 bp
 779 1415: contig of 637 bp in length
 1416 1515: gap of 100 bp
 1516 2195: contig of 680 bp in length
 2196 2295: gap of 100 bp
 2296 2983: contig of 688 bp in length
 2984 3083: gap of 100 bp
 3084 3765: contig of 666 bp in length
 3770 3869: gap of 100 bp
 3870 4673: contig of 704 bp in length
 4674 5367: gap of 100 bp
 5368 5467: gap of 100 bp

5468 6161: contig of 694 bp in length
 6162 6261: gap of 100 bp
 6262 6947: contig of 686 bp in length
 6948 7047: gap of 100 bp
 7048 7727: contig of 680 bp in length
 7728 7827: gap of 100 bp
 7828 8512: contig of 685 bp in length
 8513 8612: gap of 100 bp
 8613 9297: contig of 685 bp in length
 9298 9397: gap of 100 bp
 9398 10094: contig of 697 bp in length
 10095 10194: gap of 100 bp
 10195 10892: contig of 698 bp in length
 10893 10992: gap of 100 bp
 10993 11674: contig of 682 bp in length
 11675 11774: gap of 100 bp
 11775 12457: contig of 683 bp in length
 12458 12557: gap of 100 bp
 12558 13238: contig of 681 bp in length
 13239 13338: contig of 100 bp
 13339 14002: gap of 664 bp in length
 14003 14102: gap of 100 bp
 14103 14787: contig of 685 bp in length
 14788 14887: gap of 100 bp
 14888 15574: contig of 687 bp in length
 15575 15674: gap of 100 bp
 15675 16357: contig of 683 bp in length
 16358 16457: gap of 100 bp
 16458 17134: contig of 677 bp in length
 17135 17234: gap of 100 bp
 17235 17920: contig of 686 bp in length
 17921 18020: gap of 100 bp
 18021 18706: contig of 686 bp in length
 18707 18806: gap of 100 bp
 18807 19501: contig of 693 bp in length
 19502 19601: gap of 100 bp
 19602 20302: contig of 701 bp in length
 20303 20402: gap of 100 bp
 20403 21095: contig of 693 bp in length
 21096 21195: gap of 100 bp
 21196 21880: contig of 685 bp in length
 21881 21980: gap of 100 bp
 21981 22670: contig of 690 bp in length
 22671 22770: gap of 100 bp
 22771 23445: contig of 675 bp in length
 23446 23545: gap of 100 bp
 23546 24237: contig of 687 bp in length
 24238 24337: gap of 100 bp
 24339 25008: contig of 676 bp in length
 25009 25108: gap of 100 bp
 25109 25800: contig of 692 bp in length
 25801 25900: gap of 100 bp
 25901 26602: contig of 702 bp in length
 26603 26702: gap of 100 bp
 26703 27385: contig of 683 bp in length
 27386 27485: gap of 100 bp
 27486 28169: contig of 684 bp in length
 28170 28269: gap of 100 bp
 28270 28944: contig of 675 bp in length
 28945 29044: gap of 100 bp
 29045 29735: contig of 691 bp in length
 29736 29835: gap of 100 bp
 29836 30515: contig of 680 bp in length
 30516 30615: gap of 100 bp
 30616 31268: contig of 674 bp in length
 31269 31368: gap of 100 bp
 31369 32065: contig of 680 bp in length
 32070 32169: gap of 100 bp
 32170 32856: contig of 687 bp in length
 32857 32956: gap of 100 bp
 32957 33653: contig of 697 bp in length
 33654 33753: gap of 100 bp
 33754 34452: contig of 699 bp in length

```

* 34453 34552: gap of 100 bp
* 34553 35256: contig of 704 bp in length
* 35257 35356: gap of 100 bp
* 35357 36056: contig of 700 bp in length
* 36057 36156: gap of 100 bp
* 36157 36851: contig of 695 bp in length
* 36852 36951: gap of 100 bp
* 36952 37633: contig of 682 bp in length
* 37634 37733: gap of 100 bp
* 37734 38427: contig of 694 bp in length
* 38428 38527: gap of 100 bp
* 38528 39201: contig of 674 bp in length
* 39202 39301: gap of 100 bp
* 39302 39985: contig of 684 bp in length
* 39986 40085: gap of 100 bp
* 40086 40755: contig of 670 bp in length
* 40756 40855: gap of 100 bp
* 40856 41535: contig of 680 bp in length
* 41536 42339: contig of 704 bp in length
* 42340 42439: gap of 100 bp
* 42440 43123: contig of 684 bp in length
* 43124 43223: gap of 100 bp
* 43224 43905: contig of 682 bp in length
* 43906 44005: gap of 100 bp
* 44006 44686: contig of 681 bp in length
* 44687 44786: gap of 100 bp
* 44787 45483: contig of 697 bp in length
* 45484 45583: gap of 100 bp
* 45584 46266: contig of 683 bp in length
* 46267 46366: gap of 100 bp
* 46367 47049: contig of 683 bp in length
* 47050 47149: gap of 100 bp
* 47150 47834: contig of 685 bp in length
* 47835 47934: gap of 100 bp
* 47935 48623: contig of 689 bp in length
* 48624 48723: gap of 100 bp
* 48724 49408: contig of 685 bp in length
* 49409 49508: gap of 100 bp
* 49509 50213: contig of 705 bp in length
* 50214 50313: gap of 100 bp
* 50314 51024: contig of 711 bp in length
* 51025 51124: gap of 100 bp
* 51125 51800: contig of 676 bp in length
* 51801 51900: gap of 100 bp
* 51901 52591: contig of 691 bp in length
* 52592 52691: gap of 100 bp
* 52692 53436: contig of 745 bp in length
* 53437 53536: gap of 100 bp
* 53537 54215: contig of 679 bp in length
* 54216 54315: gap of 100 bp

```

Query Match 6.0%; Score 73.6; DB 2; Length 61326;
 Best Local Similarity 66.3%; Pred. No. 0.018; DB
 Matches 122; Conservative 0; Mismatches 59; Indels 3; Gaps 1;

```

QY 203 CCGGGTCAAGGCGCAACCTGTCAGAGTGTGTCACAGTGCAGGCGGCGTCAAGTGA 262
Db 35594 CCAAGGTCAAGGCGCAACCTGTCAGAGTGTGTCACAGTGCAGGCGGCGTCAAGTGA 35538
QY 263 CCAGCAGTGAAGTCTATGAGGCGCTTCACCTGTCATGCATCCAGAACATCACTTCTCCCTG 322
Db 35537 CCAATGCAGAGGAGCTATGAGACCTTCACTGTCATGCATCCAGAACATCACTTCTCCCTG 35478
QY 323 TCACCTTTAGAGAGCTGCGCCCTTACAGCAGCTGCGTGGCTGCGCTCCCTCCCTG 382
Db 35477 TCACCTTTGAGGAGCTGCTGAGGAGGAGTTCAGAGTGTGAGGAGCAAGGCTTCTCTAG 35418
QY 383 TCCT 386
Db 35417 TCCT 35414

```

```

RESULT 12
AF429315 125020 bp DNA linear PRI 18-JAN-2002
LOCUS Homo sapiens junctophilin 3 (JPH3) gene, partial cds.
DEFINITION AF429315
ACCESSION AF429315
VERSION AF429315.1 GI:17646244
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 125020)
Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S.,
Ingersoll-Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A.,
Potter,N.T., Ross,C.A. and Margolis,R.L.
A repeat expansion in the gene encoding junctophilin-3 is
associated with Huntington disease-like 2
Nat. Genet. 29 (4), 377-378 (2001)
PUBMED 11694876
JOURNAL MEDLINE 21583737
TITLE 2 (bases 1 to 125020)
Ingersoll-Ashworth,R.G., Ross,C.A. and Margolis,R.L.
HOLMES,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S.,
Ingersoll-Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A.,
Potter,N.T., Ross,C.A. and Margolis,R.L.
Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical
Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA
JOURNAL TITLE Location/Qualifiers
FEATURES
source 1..125020
"organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/map="16q24.3; between D16S520 and WI-12410"
/note="Isolated from a patient with Huntington's
Disease-like 2 (HDL2)"
complement(35581..35746)
/rpt_type=tandem
/rpt_unit=ctg
complement(<36507..36887)
/gene="JPH3"
/product="Junctophilin 3"
complement(<36507..36887)
/gene="JPH3"
/note="JPH3"
complement(<36507..36887)
/gene="JPH3"
/note="component of the junctional complex between plasma
membrane and endoplasmic reticulum"
/codon_start=1
/product="Junctophilin 3"
/protein_id="AAL40941.1"
/db_xref="GI:17646245"
/translation="MSSGGRRFDDGGSGYCGGMEGDKAHGHCVCIPKGGGRTGTSWS
HCFEVLGVTWPSGNTVGITWAGRRHGLGDSKGGWYKGMTWGFGRIGVRCAG
NGAKYEGTWSNLDGIGTETSDG"
BASE COUNT 29056 a 32731 c 30696 g 28283 t 4254 others
ORIGIN

```

Query Match 5.5%; Score 68; DB 9; Length 125020;
 Best Local Similarity 12.2%; Pred. No. 0.11;
 Matches 104; Conservative 370; Mismatches 377; Indels 4; Gaps 3;

```

QY 375 CCGCTGAGTCCTCTGACGCGCTGTCGCGCGCGCTATGTCAGGCGCGTCAA 434
Db 16877 MMSHAGRRMCTCYSSWMSMBMSYSYKMHMSBSGCMHBKCMTNSCCSMYKSS 16936
QY 435 CGTGCTGCTCTGTACAGAGCGGTATGGGAGGTGAGATTAACGAGCGGAAGCTTA 494
Db 16937 WMSWSMGCMCMWRGRKGMWYSRGMRSKSMRYTGSGKMMRSMCTSSCYASMCMMC 16996
QY 495 CGACGCTACGCTCTCTACAGCAGCTGCCCCGAGGAGACCGCAAGTGTGTAACCTCATCT 554
Db 16997 WSCCMBRSCCMCRSYCCCMRYCACCKCYMSSYMTWSASYSRYSRWKSMKMSRMC 17056
QY 555 AAAGCCGAGCTGAGAGCGGCGTGGGGGCTAACAGCTCTTCTGAGAGACCGGAGCTCTCT 614

```

```

Db 17057 SSSSSCKRRGGGSGMKGKGRKSYGRKTRKRSKMKMGAKWYHYRRSRMKMKMSSKRM 17116
OY 615 GCGCGCGCCCTGAGCCCTCCGCCGACCTCTGTGTAACCTGAGCCGCTGCGCGAGCTGCT 674
Db 17117 YCMTCWCKRRGACYSOCMTSRASAMCCSYCAKCKSMCYGYGMSKGYVYACSRGMS 17176
OY 675 CGTGTGCTTTTGAGACGCTCTCTCTGAGCCGCGGCTGTGACAGCAGCAGCTT-CCGAGAG 733
Db 17177 SKTCMRSTTSTGCGCCCTTTTCCCCCNANMTGGGGAACCTTTNCKNTYVYKRRGC 17236
OY 734 GCGCTCCCGGCTGTGAGCTTACCCGACGACCCATCTTCATACCTTCGAGGCGGCGAG- 792
Db 17237 AMCKYNNYNNMSRSCRAMSCTRYKSSMTSMASMYCWMCMYCYCSMRSMASRGMS 17296
OY 793 -AGCGCGACCCCGCGACCCGCGGCTCGGCTGTGCGCGACGACCCGCGAGCTGTGAC 851
Db 17297 YMKMMSMRMSCYMKCMCKMCKTCKMMSMRSMRSMASMYCYCYKSMSSRGCTRYWCM 17356
OY 852 CTGTGCTGTGTGAGCCGCGCTCCGTGACTCTCTCTCCGATTTTGGAAAGAGTGCA 911
Db 17357 SSKSCYKSYMMMRSMRSMKSMRSMKSMRSMKSMRSMKSMRSMKSMRSMKSMRSMK 912
OY 912 GCTGGCGCTGCGCGGAGAGCTCGCTACAGCCGCTGTGAGAGAGACCCGAGAGCAGCT 971
Db 17417 CMMKMGYRYRRCWMSRMRMRMRMRMRMRMRMRMRMRMRMRMRMRMRMRMRMRMR 17476
OY 972 GAGAGCGACAGAGACCCGCTGATCTTCTGAGAGCCGAGCTGTGAGGCGCGGCGCT 1031
Db 17477 CASCRCAMMSGTMYKCKASYMCKMSYTCWCTSYMSYRCTCCKMSYTSKMSWSS 17536
OY 1032 GAGCTCAGAGGAGCGCGCGGAGCGCTGAGAGGCG-CACTGTGCTGTGCGGCGCTGTTTG 1090
Db 17537 MSSYKRRKRSYSMOCISRGMSCKRCYRGASSRSMRSMRSMRSMRSMRSMRSMRSM 17596
OY 1091 GAGACCCCTCAGCTCCACCGCACCGAGTGGGTCTGCTGTGAGAGAGCCGAGAGCG 1150
Db 17597 TGMMSRSMKSYCTGRMMMTYMCCKRRSYMYTYSMSAMGRKSSMGMMMSASASR 17656
OY 1151 AAGTGAGAGCTGTGATGTGCTGCGTGGGAACTACAGTGGTCCGCGACACTTCTACTGCC 1210
Db 17657 RCKSASRSSWCSRMKMRSMKSMKSMKSMKSMKSMKSMKSMKSMKSMKSMKSMK 17716
OY 1211 TGTGTCCAGAGATG 1225
Db 17717 RGRKRSMTKSGSKG 17731

RESULT 13
AC091454_2
SEQUENCE split into 4 fragments LOCUS AC091454 Accession AC091454
Fragment Name Begin End
AC091454_0 1 110000
AC091454_1 100001 210000
AC091454_2 200001 310000
AC091454_3 300001 394326
Continuation (3 of 4) of AC091454 from base 200001 (AC091454 Mus musculus chromosome X c
Query Match 5.3%; Score 65.4; DB 2: Length 110000;
Best Local Similarity 45.6%; Pred. No 0.27;
Matches 231; Conservative 0; Mismatches 276; Indels 0; Gaps 0;

OY 355 GTGGCTGCGGCTGAGCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 414
Db 68699 GTTCCCTTCTGTCGCGTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 68758
OY 415 TATGTCAAGTCCGCTCAACAGTGTGCTGTGTACAGAGAGCGATGAGGAGTGAG 474
Db 68759 CTGCTCCCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 68818
OY 475 ATAAACGAGGAGGAGCTCTGAGAGGCTTACGTCTCTAAGAGGACTGCGCCGAGAGACCGC 534

```

```

Db 68819 CTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 68878
OY 535 AAGTGTGAACCTTATCTTAAAGCCGACGCTGAGAGCGGCTGAGGAGCTCAAGCTCTTC 594
Db 68879 CTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 68938
OY 595 CTGGAAGACGAGGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 654
Db 68939 CTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 68998
OY 655 AACCGCTGCGAGCGCTCATGTGAGTGTGAGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 714
Db 68999 CTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 69058
OY 715 AACCAACCTTCCGAGAGGCTGTGCGGCTGTGAGAGCTCAACCGGACATCTTC 774
Db 69059 CTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 69118
OY 775 ATCACTTCGAGGAGCGAGAGCGGACCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 834
Db 69119 CTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 69178
OY 835 CACCGCCACCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 861
Db 69179 CTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 69205

RESULT 14
AC105364
LOCUS AC105364 145936 bp DNA linear HTG 03-JAN-2002
DEFINITION Oryza sativa chromosome 3 clone OJ143A09, *** SHORUNTING IN
ACCESSION AC105364
VERSION AC105364.1 GI:18042282
KEYWORDS HTG; PHASE2.
SOURCE Oryza sativa.
ORGANISM Oryza sativa.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzoideae; Oryza.
REFERENCE 1 (bases 1 to 145936)
Wing, R.A., Yu, Y., Soderlund, C., Chen, M., Kim, H.-R., Rambo, T.,
Saski, C., Henry, D., Oates, R. and Simmons, J.
Rice Genomic Sequence
Unpublished
2 (bases 1 to 145936)
Wing, R.A., Yu, Y., Soderlund, C., Chen, M., Kim, H.-R., Rambo, T.,
Saski, C., Henry, D., Oates, R. and Simmons, J.
Submitted (03-JAN-2002) Clemson University Genomics Institute,
Clemson University, 100 Jordan Hall, Clemson, SC 29634 USA
NOTE: This is a 'working draft' sequence. It currently
consists of 13 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces
is believed to be correct as given, however the sizes
of the gaps between them are based on estimates that have
been provided by the submitter.
This sequence will be replaced
by the finished sequence as soon as it is available and
the accession number will be preserved.
1
7522 10885: contig of 7521 bp in length
gap of unknown length
10886 13249: contig of 3364 bp in length
gap of unknown length
13250 14250: contig of 2364 bp in length
gap of unknown length
14251 15828: contig of 1001 bp in length
gap of unknown length
15829 32112: contig of 1578 bp in length
gap of unknown length
32113 34467: contig of 16284 bp in length
gap of unknown length
34467: contig of 2355 bp in length

```


Fri Nov 8 09:52:55 2002

us-09-598-443-1.rge

Page 16

[illegible]

Search completed: November 7, 2002, 18:25:52
Job time : 2663 secs

DR P-PSDB: AAY25426.
 XX New Interleukin-1 receptor analog SIGIRR nucleic acid and proteins
 XX
 PS Claim 1: Page 68-69; 72pp; English.
 XX

CC This invention describes a novel human SIGIRR DNA, its allelic variants
 CC or species homologs which have anti-inflammatory and anti-autolymne
 CC disease activity. SIGIRR is an analog of interleukin-1 receptor. The
 CC products of the invention are used (a) as probes or primers for
 CC identifying nucleic acid that encodes proteins with SIGIRR activity; (b)
 CC to identify human chromosome 11, to map genes on this chromosome and to
 CC identify disease-related genes (particularly in the region 11p15.5 where
 CC genes are present associated with e.g. arthrogryposis multiplex
 CC congenita, breast cancer, insulin-dependent diabetes, sickle cell
 CC anaemia, bladder cancer), including detection of defective genes; (c)
 CC to study cell-signal transduction and the SIGIRR system; and (d) in
 CC gene therapy. Sense and antisense oligonucleotides can be used to inhibit
 CC expression of the SIGIRR gene. The proteins of the invention are used:
 CC (1) to study cellular processes (immune regulation, proliferation, death,
 CC migration, interaction with other cells and inflammation); (ii) to
 CC identify and purify proteins that associate with SIGIRR ligands and
 CC receptors, and to measure their biological activity; (iii) in screening
 CC for, and rational design of, potential inhibitors of activity; (iv)
 CC therapeutically against diseases mediated by SIGIRR polypeptide
 CC counter-structures; (v) as molecular weight (m.w.) markers in
 CC electrophoresis; (vi) for determining isoelectric points of unknown
 CC proteins; (vii) as controls for determining the extent of protein
 CC fragmentation (e.g. to aid characterization of protein structures by
 CC mass spectrometry); (viii) for generation of antibodies (Ab); (ix) to
 CC deliver diagnostic or therapeutic agents to cells that express SIGIRR
 CC binding molecules.
 XX

SO Sequence 1233 BP; 203 A; 408 C; 380 G; 242 T; 0 other;

Query Match 100.0%; Score 1233; DB 20; Length 1233;
 .Best Local Similarity 100.0%; Pred. No. 4.3e-239;
 Matches 1233; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCCAGGAGTGTGTGATAGGCCCCCTGACTTCTCCCGCTGAAGACAGAGTGTG 60
 DB 1 ATCCAGGAGTGTGTGATAGGCCCCCTGACTTCTCCCGCTGAAGACAGAGTGTG 60
 QY 61 AGGCTGCTTGGGAGCTCAAGTGTGAACTGCACAGGCTTGGTGTCTCTGAGGCC 120
 DB 61 AGGCTGCTTGGGAGCTCAAGTGTGAACTGCACAGGCTTGGTGTCTCTGAGGCC 120
 QY 121 CACTGCTCCCTGCTTCACTCACTGAGTGAAGAGGAGGCTTCAATTGGAGGAGC 180
 DB 121 CACTGCTCCCTGCTTCACTCACTGAGTGAAGAGGAGGCTTCAATTGGAGGAGC 180
 QY 181 CACTAGAGCTCCAGAGTACTCTGGGCAAGGCCAAGCTGTAGAGAGTGTGTGTC 240
 DB 181 CACTAGAGCTCCAGAGTACTCTGGGCAAGGCCAAGCTGTAGAGAGTGTGTGTC 240
 QY 241 AGTGTCTTGGGGGTCAAGTGAACAGCACTGAAGTCTTGGGGCTTACTGCTCATC 300
 DB 241 AGTGTCTTGGGGGTCAAGTGAACAGCACTGAAGTCTTGGGGCTTACTGCTCATC 300
 QY 301 CAGACATCAGCTTCTCTCTCTTCACTCTAGAGAGCTGAGCTTACAGAGCAGTGGCT 360
 DB 301 CAGACATCAGCTTCTCTCTCTTCACTCTAGAGAGCTGAGCTTACAGAGCAGTGGCT 360
 QY 361 GCGGTCTGAGCTTCT 420
 DB 361 GCGGTCTGAGCTTCT 420
 QY 421 AAGTGCCTGTCACAGCTGTCTGTGTACAGAGAGGAGGAGTGGAGGATGAATTAAC 480
 DB 421 AAGTGCCTGTCACAGCTGTCTGTGTACAGAGAGGAGTGGAGGATGAATTAAC 480
 QY 481 GAGGGAGAGCTCTAGAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
 DB 481 GAGGGAGAGCTCTAGAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540

DB 481 GAGGGAGAGCTCTAGAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
 QY 541 GTGAATTCATCTTAAGCCAGCTGAGAGCGGCTGAGGAGCTTCTCTCTGAG 600
 DB 541 GTGAATTCATCTTAAGCCAGCTGAGAGCGGCTGAGGAGCTTCTCTCTGAG 600
 QY 601 GACCGGAGCT 660
 DB 601 GACCGGAGCT 660
 QY 661 TCCGAGAGCT 720
 DB 661 TCCGAGAGCT 720
 QY 721 ACCTTCGAGAGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
 DB 721 ACCTTCGAGAGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
 QY 781 TTGAGAGGAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
 DB 781 TTGAGAGGAG 840
 QY 841 CACCTGAGTACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
 DB 841 CACCTGAGTACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
 QY 901 AAGAGAGTACAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
 DB 901 AAGAGAGTACAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
 QY 961 CAG 1020
 DB 961 CAG 1020
 QY 1021 GCGCGGAG 1080
 DB 1021 GCGCGGAG 1080
 QY 1081 CCTGTTTTGGAG 1140
 DB 1081 CCTGTTTTGGAG 1140
 QY 1141 CGAG 1200
 DB 1141 CGAG 1200
 QY 1201 TTCTACTGCTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1260
 DB 1201 TTCTACTGCTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1260

RESULT 2
 AAF27859
 ID AAF27859 standard; cDNA: 1659 BP.
 XX
 AC AAF27859;
 XX
 DT 30-MAR-2001 (first entry)
 XX
 DE Human NOV9 cDNA.
 XX
 KW Human; NOV9; antiinflammatory; cyostatic; neuroprotective;
 KW cerebrioprotective; immunomodulator; vulnery; vasolropic; gene therapy;
 KW hyperplasia; tumour; restenosis; psoriasis; Dupuytren's contracture;
 KW diabetes; rheumatoid arthritis; cerebral oedema; Alzheimer's disease; ss.
 OS Homo sapiens.
 XX
 PN WO200075321-A2.
 XX
 PD 14-DEC-2000.
 XX
 PF 01-JUN-2000; 2000WO-US15303.

XX 03-JUN-1999: 99US-0137322.
PR 16-MAR-2000: 2000US-0189810.
PR 22-MAR-2000: 2000US-0191158.
PR 30-MAR-2000: 2000US-0193086.
PR 31-MAY-2000: 2000US-0137322.
XX
XX (CURA-) CURAGEN CORP.
PI Shinkets RA, Fernandes E, Herrman J, Vernet C;
XX
XX WPI: 2001-102403/11.
DR P-PSDB: AAB61137.
XX
XX New NOXV polypeptides and polynucleotides, useful in gene therapy, as a
PT diagnostic marker, protein therapeutic, antibody or small molecule drug
PT target for treating immune, proliferative and metabolic diseases and
PT wound healing
XX
XX Laim 8: Page 36-38; 194P; English.
XX
XX The present sequence encodes a new isolated polypeptide (NOXV). The NOXV
CC polypeptides, NOXV nucleic acids, and anti-NOXV antibodies are useful for
CC treating or preventing NOXV-associated disorders. They are also useful
CC for determining the presence of or a predisposition to a disease
CC associated with altered levels of the NOXV polypeptide or nucleic acid.
CC These NOXV-associated disorders include hyperplasias, tumours,
CC rheumatoid arthritis, Dupuytren's contracture, diabetic complications,
CC oedema, senile dementia or Alzheimer's disease. The NOXV polynucleotides
CC are especially useful in gene therapy. Specifically, NOXV is useful as
CC a diagnostic marker or prognostic marker, protein therapeutic and
CC antibody target or small molecule drug target to treat disorders in the
CC immune response pathway, thyroid and metabolic diseases, bone metabolic
CC disorders, diseases of the pancreas (e.g. diabetes or digestive
CC disorders), proliferative diseases, or tissue regeneration and
CC development (e.g. wound healing or treatment of burns).
XX
XX Sequence 1659 BP; 296 A; 551 C; 504 G; 308 T; 0 other:
SQ
Query Match 99.9%; Score 1231.4; DB 22; Length 1659;
Best Local Similarity 99.9%; Pred. No. 9.4e-239;
Matches 1332; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 664 AAGTGCCTCTCAACGTCTGCTCTGTACACAGAGCGCTATGGGAGGTGAGATTAAC 723
Qy 481 GACGGGAACCTCTACGACCCCTACGCTCTCTACAGAGACTGCCCGAGAGCGCAAGTTC 540
Db 724 GACGGGAACCTCTACGACCCCTACGCTCTCTACAGAGACTGCCCGAGAGCGCAAGTTC 783
Qy 541 GTGAACCTTATCTTAAGCGGAGCTGAGCGGCGCTGGGGCTACAAAGCTTCTCCGAGC 600
Db 784 GTGAACCTTATCTTAAGCGGAGCTGAGCGGCGCTGGGGCTACAAAGCTTCTCCGAGC 843
Qy 601 GACCGGACCTCTGCGCGCGCTGAGCCCTCCGCGACCTTGGTGAACCTGAGCCGC 660
Db 844 GACCGGACCTCTGCGCGCGCTGAGCCCTCCGCGACCTTGGTGAACCTGAGCCGC 903
Qy 661 TGCCGACGCTCATGCTGCTGCTTTTGAGAGCGCTTCTTACGCCGGGCTGTGTGAGCCAC 720
Db 904 TGCCGACGCTCATGCTGCTGCTTTTGAGAGCGCTTCTTACGCCGGGCTGTGTGAGCCAC 963
Qy 721 AGCTTCGGGAGGCGCTGTGCGGCTGTGAGCTCACCGCAGACCCATCTTCATCACC 780
Db 964 AGCTTCGGGAGGCGCTGTGCGGCTGTGAGCTCACCGCAGACCCATCTTCATCACC 1023
Qy 781 TTGAGAGGCGCAGAGCGCGACCCCGGCAACCGGCGCTCCGCTGTGCGCAGCACCGC 840
Db 1024 TTGAGAGGCGCAGAGCGCGACCCCGGCAACCGGCGCTCCGCTGTGCGCAGCACCGC 1083
Qy 841 CACCTGTACCTTGTGCTGTGTGAGAGCGCGCTCCGTACATCTTCTCCGATTTTGG 900
Db 1084 CACCTGTACCTTGTGCTGTGTGAGAGCGCGCTCCGTACATCTTCTCCGATTTTGG 1143
Qy 901 AAAGAAGTACACCTGCGCTGCGCGGAGGTCGCTACAGGCGGCTGGAAGAGACCC 960
Db 1144 AAAGAAGTACACCTGCGCTGCGCGGAGGTCGCTACAGGCGGCTGGAAGAGACCC 1203
Qy 961 CAGACGACCTCGAGAGCAGACAGAACCCCATGTGATTTTTCAGAGCGCGAGTCCCTGAG 1020
Db 1204 CAGACGACCTCGAGAGCAGACAGAACCCCATGTGATTTTTCAGAGCGCGAGTCCCTGAG 1263
Qy 1021 GGC CGCGCGCTGACCTCAAGGTGAGCCCGGACCTCGAGAGCGGAGTGTGCGGGGG 1080
Db 1264 GGC CGCGCGCTGACCTCAAGGTGAGCCCGGACCTCGAGAGCGGAGTGTGCGGGGG 1323
Qy 1081 CCTGTTTGTGAGAGCCTACAGCTCCACCGCACACAGTGGGGTCTCGGTGGAGAGAGC 1140
Db 1324 CCTGTTTGTGAGAGCCTACAGCTCCACCGCACACAGTGGGGTCTCGGTGGAGAGAGC 1383
Qy 1141 CGAGAGCGGAAGTGAAGCTCTCGATCTCGGCTCGGAACCTACAGTCCCGCACAGAC 1200
Db 1384 CGAGAGCGGAAGTGAAGCTCTCGATCTCGGCTCGGAACCTACAGTCCCGCACAGAC 1443
Qy 1201 TTCTACTGCTGTGTCTCCAGAGATGATATGATAG 1233
Db 1444 TTCTACTGCTGTGTCTCCAGAGATGATATGATAG 1476
RESULT 3
AAS27325
ID AAS27325 standard; cDNA; 1428 BP.
XX
XX AAS27325;
XX
XX 07-NOV-2001 (first entry)
XX
XX cDNA encoding novel signal transduction pathway protein, Seq ID 360.
XX
XX Neuroprotective; cytoskeletal; dermatological; immunosuppressive; tumour;
KW anti-inflammatory; anti-HIV; antibacterial; anti-inflammatory condition;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;

KM cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS; ss;
XX acquired immune deficiency syndrome.
OS Homo sapiens.
XX WO200154733-A1.
XX PD 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US01312.
XX PF 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 03-JUN-2000; 2000US-0214886.
PR 05-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232081.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.

PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 25-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.

PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Rosen CA, Barash SC, Ruben SK;
 PI WPI: 2001-465460/50.
 DR P-PSDB: AAU17408.
 XX
 PR Novel polypeptides useful for diagnosing, treating, preventing and/or
 PT prognosing disorders related to the proteins, including cancers, immune
 PT disorders and neuronal disorders
 PS
 PS Claim 1; SEQ ID No 360; 880bp; English.

CC The invention relates to novel isolated polypeptides (I), and
 CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
 CC diagnosing, preventing and treating diseases including immune system
 CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
 CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
 CC transplant rejection and graft versus host disease, infectious diseases
 CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
 CC other blood-related disorders (sickle cell anaemia), myeloproliferative
 CC disorders, primary haematopoietic disorders, hyperproliferative
 CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
 CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
 CC disorders (e.g. glomerulonephritis), cardiovascular disorders
 CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
 CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
 CC Addison's disease), reproductive system disorders, gastrointestinal
 CC disorder (inflammatory disorders), liver disorders (cirrhosis),
 CC as stimulators of B-cell responsiveness to pathogens, activators of
 CC T-cells, to induce higher affinity antibodies, and as a means to induce
 CC tumour proliferation in pathologies e.g. acquired immune deficiency
 CC syndrome (AIDS). AA52676-AA52785 represent novel signal transduction
 CC pathway protein coding sequences and PCR primers of the invention.
 XX

Query Match 83.5%; Score 1029.4; DB 22; Length 1428;
 Best Local Similarity 99.9%; Pred. No. 4e-198;
 Matches 1030; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 203 CCTGGTCAAGGCAACCTGTGAGAGTGTGTCAGTGGCCGAGGATCAAGTGA 262
 DB 196 CCAGGGTCAAGGCCAACCCTGTAGAGTCTTGTGTCCAGTGTCTGGGATCAAGCTGA 255
 QY 263 CCAGCACTGAACTATGAGGGGCTTCACTGCTCCATCCAGAACATCAGTCTCTCTCT 322
 DB 256 CCAGCACTGAACTATGAGGGGCTTCACTGCTCCATCCAGAACATCAGTCTCTCTCT 315
 QY 323 TCACCTTTCAAGAGTGTGGCCCTTACAGCCACGTGGCTGGCTGCTCTCTCTCTG 382
 DB 316 TCACCTTTCAAGAGTGTGGCCCTTACAGCCACGTGGCTGGCTGCTCTCTCTCTG 375
 QY 383 TCCTGCTGGCCCTTGTGCTGGCCGCTCTGTATGTCAAGTGGCGTCAAGGTCGTCG 442
 DB 376 TCCTGCTGGCCCTTGTGCTGGCCGCTCTGTATGTCAAGTGGCGTCAAGGTCGTCG 435
 QY 443 TCTGTACCAAGACGCGTATGGGAGGTGAGATAAAGCAGCGGAAGCTCTACGAGCCCT 502
 DB 436 TCTGTACCAAGACGCGTATGGGAGGTGAGATAAAGCAGCGGAAGCTCTACGAGCCCT 495
 QY 503 AGCTCTCTCAACGCACTCCCCGAGAGACGCAAGTTCGTGAATTCATCTTAAAGCCGC 562
 DB 496 AGCTCTCTCAACGCACTCCCCGAGAGACGCAAGTTCGTGAATTCATCTTAAAGCCGC 555
 QY 563 AGCTGAGAGGCGGTGCGGAGCTCAAGCTCTTCTGAGACGACCGGACCTCGCGCGGCG 622
 DB 556 AGCTGAGAGGCGGTGCGGAGCTCAAGCTCTTCTGAGACGACCGGACCTCGCGCGGCG 615
 QY 623 CTGAGCCCTCGCGGACCTCTTGTGAACCTGAGCCGCTGCGAGGCGCTCATCTGTGTGTC 682

DB 616 CTGAGCCCTCCGCCGACCTTGTGTAACCTGAGCGGTGCGGACGCTCATCTGTGTGTC 675
 QY 683 TTTGGAGGCGCTTCCGTGAGACCGGCGCTGTGTGACGACGACGCTTCCGGAGGCGCTGTGTC 742
 DB 676 TTTGGAGGCGCTTCCGTGAGACCGGCGCTGTGTGACGACGACGCTTCCGGAGGCGCTGTGTC 735
 QY 743 GGTGCTGAGGCTCAACCCGACGACCCATCTTCATCAGTTCAGAGGCGCAGAGCGGAC 802
 DB 736 GGTGCTGAGGCTCAACCCGACGACCCATCTTCATCAGTTCAGAGGCGCAGAGCGGAC 795
 QY 803 CCGCGACCCGCGGCTCGGCTGTGTGAGCGGACGACCGGACCTGTGTGTGTGTGTGTGT 862
 DB 796 CCGCGACCCGCGGCTCGGCTGTGTGAGCGGACGACCGGACCTGTGTGTGTGTGTGTGT 855
 QY 863 GGAGGCGCGGCTCGGCTGTGTGAGCGGACGACCGGACCTGTGTGTGTGTGTGTGTGTGT 922
 DB 856 GGAGGCGCGGCTCGGCTGTGTGAGCGGACGACCGGACCTGTGTGTGTGTGTGTGTGTGT 915
 QY 923 CCGGAGAGGTGCGGTACAGGCGGCTGGAAGAGACCCGACGACGCTGACGAGAGACA 982
 DB 916 CCGGAGAGGTGCGGTACAGGCGGCTGGAAGAGACCCGACGACGCTGACGAGAGACA 975
 QY 983 AGAGCCCATGCTGATTTCTGAGGCGGAGTCTGAGGCGCGGCGCTGTGTGTGTGTGTGT 1042
 DB 976 AGAGCCCATGCTGATTTCTGAGGCGGAGTCTGAGGCGCGGCGCTGTGTGTGTGTGTGT 1035
 QY 1043 TGGACCCGACCTGAGGCGGACGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1102
 DB 1036 TGGACCCGACCTGAGGCGGACGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1095
 QY 1103 CTCACCCGACACGACGATGGGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1162
 DB 1096 CTCACCCGACACGACGATGGGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1155
 QY 1163 CGGATCTGCGCTGCGGAACCTACAGTGTGCGGACGACGACCTGTGTGTGTGTGTGT 1222
 DB 1156 CGGATCTGCGCTGCGGAACCTACAGTGTGCGGACGACGACCTGTGTGTGTGTGTGT 1215
 QY 1223 ATGATATGTAG 1233
 DB 1216 ATGATATGTAG 1226

RESULT 4
 AA163958
 ID AA163958 standard; cDNA; 1428 BP.

XX AA163958;

DT 22-OCT-2001 (first entry)

XX Human polynucleotide SEQ ID NO 166.

XX Human; antihypertensive; antirheumatic; antiproliferative; vasotropic;
 KW cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;
 KW fungicide; ophthalmological; cytostatic; immunosuppressive; nootropic;
 KW neuroprotective; antiallergic; hepatotropic; antidiabetic;
 KW antiinflammatory; antilucifer; vulnary; anticonvulsant; antibacterial;
 KW cardiovascular disorder; gene therapy; cancer; immune disorder;
 KW cardiovascular disease; neurological disease; infection; human; ss.

XX Homo sapiens.

OS WO200155308-A2.

PN 02-AUG-2001.

PD 17-JAN-2001; 2001WO-0501309.

PF 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0225759.
PR 22-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229345.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0232597.
PR 14-SEP-2000; 2000US-0232597.
PR 14-SEP-2000; 2000US-0232598.
PR 14-SEP-2000; 2000US-0232599.
PR 14-SEP-2000; 2000US-0232599.
PR 14-SEP-2000; 2000US-0232601.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234597.
PR 25-SEP-2000; 2000US-0234598.
PR 26-SEP-2000; 2000US-0234584.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236602.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR XX

PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0244826.
PR 01-NOV-2000; 2000US-0244826.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
PR XX
PR XX (HUMA-) HUMAN GENOME SCI INC.
PR XX Rosen CA, Barash SC, Ruben SM;
PR XX WPI; 2001-488781/53.
PR XX P-PSDB; AAM43652.
PR XX
PR XX New isolated nucleic acids and polypeptides, useful for diagnosing,
PR XX treating and/or preventing human diseases and disorders -
PR XX
PR XX Claim 1; SEQ ID NO 166; 664bp + Sequence listing; English.
PR XX

CC The invention relates to human polynucleotides (AA163803-AA164012) and
CC the encoded proteins (AA44497-AA44360) useful for preventing, treating
CC or ameliorating medical conditions e.g. by protein or gene therapy. The
CC genes were isolated from a range of human tissues disclosed in the
CC specification. The nucleic acids, proteins, antibodies and (ant)agonists
CC are useful in the diagnosis, treatment and prevention of: (a) cancer,
CC e.g. breast and ovarian cancer and other cancers of the adrenal gland,
CC bone, bone marrow, breast, gastrointestinal tract, liver, lung, or
CC urogenital; (b) immune disorders e.g. Addison's disease, allergies,
CC autoimmune hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,
CC Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pcl_sequences.
XX

Sequence 1428 BP; 288 A; 462 C; 426 G; 250 T; 2 other:
Query Match 83.5%; Score 1029.4; DB 22; Length 1428;
Best Local Similarity 99.9%; Pred. No. 4e-198;
Matches 1030; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 203 CTTGGGTCAGAGCCAACTGTCAGAGTCTTGTCCAGTGTCCGGGGGTCAACGTGA 262
DB 196 CCAGGGTCAGAGCCAACTGTCAGAGTCTTGTCCAGTGTCCGGGGGTCAACGTGA 255
OY 263 CCAGCAGTGAAGTCTATGAGGGGCTTCCAGTGTCCAGTGTCCGGGGGTCAACGTGA 322
DB 256 CCAGCAGTGAAGTCTATGAGGGGCTTCCAGTGTCCAGTGTCCGGGGGTCAACGTGA 315
OY 323 TCACCTTCAGAGAGTGGGCTTACAGAGTGTGGTGGTGGTGGTGGTGGTGGTGGTGG 382
DB 316 TCACCTTCAGAGAGTGGGCTTACAGAGTGTGGTGGTGGTGGTGGTGGTGGTGGTGG 375
OY 383 TCTCTGCTGGGCTTGTGGTGGGCGGCTTGTATGTCAGAGCGGCTTGTATGTCAGAGCG 442
DB 376 TCTCTGCTGGGCTTGTGGTGGGCGGCTTGTATGTCAGAGCGGCTTGTATGTCAGAGCG 435
OY 443 TCTGTATCCAGAGCGGCTTGTATGAGGGAGTGAAGAGCGGAGCTTGTATGAGCGCT 502
DB 436 TCTGTATCCAGAGCGGCTTGTATGAGGGAGTGAAGAGCGGAGCTTGTATGAGCGCT 495
OY 503 ACGTCTCTACAGAGCGGCTTGTATGAGGGAGCGGAGTGTGTATGTCAGTGAAGCGCGC 562
DB 496 ACGTCTCTACAGAGCGGCTTGTATGAGGGAGCGGAGTGTGTATGTCAGTGAAGCGCGC 555
OY 563 ACGTCTCTACAGAGCGGCTTGTATGAGGGAGCGGAGTGTGTATGTCAGTGAAGCGCGC 622
DB 556 ACGTCTCTACAGAGCGGCTTGTATGAGGGAGCGGAGTGTGTATGTCAGTGAAGCGCGC 615
OY 623 CTGAGCCCTCCGCGAGACTTGTGTAACCTGAGCGGCTTGTGAGCGGCTTGTGAGTGC 682
DB 616 CTGAGCCCTCCGCGCGAGACTTGTGTAACCTGAGCGGCTTGTGAGCGGCTTGTGAGTGC 675
OY 683 TTTGGAGAGCTTCTGTAGCGGGGCTGTGAGCAGACAGTCTTGGGAGGGGCTGTGCG 742
DB 676 TTTGGAGAGCTTCTGTAGCGGGGCTGTGAGCAGACAGTCTTGGGAGGGGCTGTGCG 735
OY 743 GAGTCTGTGAGTCAACCGGAGAGCCATCTTATCAGTCTTGTGAGGGGCTGTGAGCG 802
DB 736 GAGTCTGTGAGTCAACCGGAGAGCCATCTTATCAGTCTTGTGAGGGGCTGTGAGCG 795
OY 803 CCGGCGACCCGGGCGCTCGGCTGTGCGGAGACCGGCGGCTGTGAGCGTGTGCTCT 862
DB 796 CCGGCGACCCGGGCGCTCGGCTGTGCGGAGACCGGCGGCTGTGAGCGTGTGCTCT 855
OY 863 GGAGGCGCGGCGCTCGTGTGAGTCTTGTGCGGAGTGTGGAAGAGTGAAGCGGGGCTGC 922
DB 856 GGAGGCGCGGCGCTCGTGTGAGTCTTGTGCGGAGTGTGGAAGAGTGAAGCGGGGCTGC 915

OY 923 CCGGAGAGTGGCGGTACAGAGCGCGGTGGAAGAGACCCCGAGCGAGTGCAGAGGACA 982
DB 916 CCGGAGAGTGGCGGTACAGAGCGCGGTGGAAGAGACCCCGAGCGAGTGCAGAGGACA 975
OY 983 AGGACCCCATGCTGATTTCTTGCAGAGCGGAGTCCCTGAGGGCGCGGCTGTGACTAGAG 1042
DB 976 AGGACCCCATGCTGATTTCTTGCAGAGCGCGGAGTCCCTGAGGGCGCGGCTGTGACTAGAG 1035
OY 1043 TCGACCCCGGAGCTGAGGGGCGAGCTGGGTGTCCGGGGGCTTTTGGAGAGCCATCAG 1102
DB 1036 TCGACCCCGGAGCTGAGGGGCGAGCTGGGTGTCCGGGGGCTTTTGGAGAGCCATCAG 1095
OY 1103 CTCACCCCGGAGCTGAGGGGCTTCTGCTGGGAGAGAGCCGAGCAGAGTGAAGTGTCT 1162
DB 1096 CTCACCCCGGAGCTGAGGGGCTTCTGCTGGGAGAGAGCCGAGCAGAGTGAAGTGTCT 1155
OY 1163 CGGATCTGGGCTGGGGAAGTCACTGAGTCCCGGAGAGACTTACTGCTGTGTCCAGAG 1222
DB 1156 CGGATCTGGGCTGGGGAAGTCACTGAGTCCCGGAGAGACTTACTGCTGTGTCCAGAG 1215
OY 1223 ATGATATGTAG 1233
DB 1216 ATGATATGTAG 1226

RESULT 5
AA163889
ID AA163889 standard; cDNA; 1442 BP.
XX
XX AA163889;
AC XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 97.
XX
XX Human; antihypertensive; antipneumatic; antiproliferative; vasotropic;
KW cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;
KW fungicide; ophthalmological; cytostatic; immunosuppressive; nootropic;
KW neuroprotective; antiallergic; hepatotropic; antidiabetic;
KW antinflammatory; antitumor; vulnary; anticonvulsant; antibacterial;
KW cardiovascular disorder; gene therapy; cancer; immune disorder;
KW cardiovascular disorder; neurological disease; infection; human; ss.
OS Homo sapiens.
XX
XX
XX WO200155308-A2.
PD 02-AUG-2001.
XX
XX
XX 17-JAN-2001; 2001WO-US01309.
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198122.
PR 19-MAY-2000; 2000US-0205515.
PR 28-JUN-2000; 2000US-0209467.
PR 07-JUN-2000; 2000US-0215135.
PR 30-JUN-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.

Sequence 1442 BP; 297 A; 465 C; 430 G; 250 T; 0 other:

Query Match 83.5%; Score 1029.4; DB 22; Length 1442;
Best Local Similarity 99.9%; Pred. No. 4e-198;
Matches 1030; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 203 CCTGGGTCAAGGCCAACCTGTTCAGAGTGTGTGTCCAGTGTCTGGGGGTCAACGTGA 262
    |||||||
DB 204 CCAGGGTCAAGGCCAACCTGTTCAGAGTGTGTGTCCAGTGTCTGGGGGTCAACGTGA 263
QY 263 CCAGACACTGAAGTCTATGAGGGGCTTCACCTGTCTCATCCAGAACTCAGCTTCTCTCT 322
    |||||||
DB 264 CCAGACACTGAAGTCTATGAGGGGCTTCACCTGTCTCATCCAGAACTCAGCTTCTCTCT 323
QY 323 TCACCTCTCAGAGACTGGGCCCTTACAGCCAGGTGGTGGGCTGCTCCCTCCCTGG 382
    |||||||
DB 324 TCACCTCTCAGAGACTGGGCCCTTACAGCCAGGTGGTGGGCTGCTCCCTCCCTGG 383
QY 383 TCCCTGCTGGCCCTGCTGCTGGGCCGCCCTGCTCTATGTCAAGTGCCTGTCAACGTCTGC 442
    |||||||
DB 384 TCCCTGCTGGCCCTGCTGCTGGGCCGCCCTGCTCTATGTCAAGTGCCTGTCAACGTCTGC 443
QY 443 TCTGTGTACCGAGACCGCGATGAGGGAGTGGAGATTAACGAGCGGAAGCTTACGACGCT 502
    |||||||
DB 444 TCTGTGTACCGAGACCGCGATGAGGGAGTGGAGATTAACGAGCGGAAGCTTACGACGCT 503
QY 503 ACGTCTCTACAGCGACGCTGCCCCGAGAGACCGCAATGTGTGACTTCACTCTTAAGCCGC 562
    |||||||
DB 504 ACGTCTCTACAGCGACGCTGCCCCGAGAGACCGCAATGTGTGACTTCACTCTTAAGCCGC 563
QY 563 AGCTGGAGCGGCGTGCAGAGCTCTTCTGTGAGACGACGCGGACCTCTGCGCGCGC 622
    |||||||
DB 564 AGCTGGAGCGGCGTGCAGAGCTCTTCTGTGAGACGACGCGGACCTCTGCGCGCGC 623
QY 623 CTGAGCCCTCGCGCGACCTCTTGTGTAACCTGAGCCGCTGCCGACCTCATCTGTGTGC 682
    |||||||
DB 624 CTGAGCCCTCGCGCGACCTCTTGTGTAACCTGAGCCGCTGCCGACCTCATCTGTGTGC 683
QY 683 TTTGGAGAGCCCTTCTGTGAGCGGGGCTGTGAGCACAGCTTCGGGAGGGGCTGTGCC 742
    |||||||
DB 684 TTTGGAGAGCCCTTCTGTGAGCGGGGCTGTGAGCACAGCTTCGGGAGGGGCTGTGCC 743
QY 743 GGCCTCTGAGAGCTCAACCCGAGAACCATCTTCATCACCCTTCGAGGGGCGAGAGCGCGACC 802
    |||||||
DB 744 GGCCTCTGAGAGCTCAACCCGAGAACCATCTTCATCACCCTTCGAGGGGCGAGAGCGCGACC 803
QY 803 CCGCGCACCCGGCGCTCCGCTGCTGCGCGACAGACCGCCACCTGTGACCTTGTCTCT 862
    |||||||
DB 804 CCGCGCACCCGGCGCTCCGCTGCTGCGCGACAGCACCGCCACCTGTGACCTTGTCTCT 863
QY 863 GGAGGCGCGGCTCGTGAATCTTCTCCGANTTTTGAAGAAAGTGAAGTGAAGTGGCGTGC 922
    |||||||
DB 864 GGAGGCGCGGCTCGTGAATCTTCTCCGANTTTTGAAGAAAGTGAAGTGAAGTGGCGTGC 923
QY 923 CGCGGAAGTGTGAGTCAAGGCGGTGGAAGAGAACCCCGACAGCAGTGTGAGAGAGCA 982
    |||||||
DB 924 CGCGGAAGTGTGAGTCAAGGCGGTGGAAGAGAACCCCGACAGCAGTGTGAGAGAGCA 983
QY 983 AGGACCCCATGCTGATTTCTTGAAGGCCAGTCTCTGAGGGCCGGGCTCGCATAGAG 1042
    |||||||
DB 984 AGGACCCCATGCTGATTTCTTGAAGGCCAGTCTCTGAGGGCCGGGCTCGCATAGAG 1043
QY 1043 TGGACCCCGAGACCTGAGGGGCGACGTGGGTGTCCGGGGGCTGTTTGGAGAGCCATAG 1102
    |||||||
DB 1044 TGGACCCCGAGACCTGAGGGGCGACGTGGGTGTCCGGGGGCTGTTTGGAGAGCCATAG 1103
QY 1103 CTCCACCGCACACACAGTGGGTCTTCTGAGAGAGAGCCGAGAGCAGCAAGTGAAGTCT 1162
    |||||||
DB 1104 CTCCACCGCACACACAGTGGGTCTTCTGAGAGAGAGCCGAGAGCAGCAAGTGAAGTCT 1163
QY 1163 CGGATCTTGGGCTCGGGAACATCAATGATGCGCGACAGACTTCTACTGCTGTGTCTCAAG 1222
    |||||||
DB 1164 CGGATCTTGGGCTCGGGAACATCAATGATGCGCGACAGACTTCTACTGCTGTGTCTCAAG 1223
```

QY 1223 ATGATATGTAG 1233
|||||
DB 1224 ATGATATGTAG 1234

RESULT 6
AAFI8292
ID AAFI8292 standard; DNA; 1296 BP.

AAFI8292;

14-MAR-2001 (first entry)

Lung cancer associated polynucleotide sequence SEQ ID 311.

Human; lung cancer associated protein; neuroprotective; cytoskeletal;
cardioactive; immunomodulatory; muscular active; vulnary;
gastrointestinal; nephrotropic; antiinfective; gynecological;
antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
proliferative disorder; wound healing; infectious disease; ds.

Homo sapiens.

MO20005180-A2.

21-SEP-2000.

08-MAR-2000; 2000WO-US05918.

12-MAR-1999; 99US-0124270.

(HUMA-) HUMAN GENOME SCI INC.

(ROSE/) ROSEN C A.

Ruben SM;

WPI: 2000-587514/55.

P-PSDB; AAB58416.

Lung cancer associated gene sequences, referred to as lung cancer
antigens, useful for treatment, prevention, and diagnosis of disorders
such as lung cancer -
Claim 1; Page 769-770; 1425bp; English.

Polynucleotide sequences AAFI7982 - AAFI8424 encode human lung cancer
associated proteins represented in AAB58106 - AAB58548. Lung cancer
associated proteins and polynucleotide sequences, their agonists, and
antagonists may have neuroprotective; cytostatic; cardioactive;
immunomodulatory; muscular active general; vulnary; gastrointestinal
general; nephrotropic; antiinfective; gynecological; or antibacterial
activity. The invention also includes antibodies specific for the
protein or polynucleotide sequences. The lung cancer associated
polynucleotide sequences may be used for detection of lung cancer,
chromosome identification, as chromosome markers, and for numerous other
diagnostic or research purposes. The proteins may be used to treat
disorders such as neural, immune, muscular, reproductive,
gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
disorders. The proteins may also be used in the treatment of wounds and
infectious diseases. Polynucleotide sequences AAFI8425 - AAFI8433 and
peptide AAB58549 are used in the course of the invention for the
identification and characterisation of the polynucleotide and protein
sequences.

Sequence 1296 BP; 242 A; 431 C; 394 G; 227 T; 2 other:

Query Match 83.1%; Score 1025; DB 21; Length 1296;
Best Local Similarity 99.7%; Pred. No. 3e-197;
Matches 1025; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
QY 206 GGGTCAAGGCCAACCTGTTCAGAGTGTGTGTCCAGTGTCTGGGGGTCAACGTGACCA 265
    |||||||
```


Db 67 GGGTCAAGGCCAACCTCTCAGAGGTGCTGTGTACAGTCTCCTGGGGGTCAACGTGACCA 126
 QY 266 GCACCTAAGTCTATGAGGCGCTTACCTGCTCATCCAGAACATCAGCTTCTCTCTTCA 325
 Db 127 GCACTGAAGTCTATGAGGCGCTTACCTGCTCATCCAGAACATCAGCTTCTCTCTTCA 186
 QY 326 CTCTTCAGAGAGCTGGCCCTACAAAGCCAGTGGCTGGTGGCTCCCTCTGAGCC 385
 Db 187 CTCTTCAGAGAGCTGGCCCTACAAAGCCAGTGGCTGGTGGCTCCCTCTGAGCC 246
 QY 386 TGTGAGCCCTGCTGGCCGCTGCTCATGTAAGTGGCTGTCACAGTGGCTCT 445
 Db 247 TGTGAGCCCTGCTGGCCGCTGCTCATGTAAGTGGCTGTCACAGTGGCTCT 306
 QY 446 GGTACCAAGAGCGGTATGAGGAGGTGAGATAAGCAGGGAAGCTTACAGAGCTTACG 505
 Db 307 GGTACCAAGAGCGGTATGAGGAGGTGAGATAAGCAGGGAAGCTTACAGAGCTTACG 366
 QY 506 TCTCTCAGAGAGTGGCCCTGAGAGACCCAGTCTGTGTAATTCATCTTAAAGCCGACG 565
 Db 367 TCTCTCAGAGAGTGGCCCTGAGAGACCCAGTCTGTGTAATTCATCTTAAAGCCGACG 426
 QY 566 TGAAGGAGCGTGGGGCTACAGCTCTCTCTGAGAGACCGGAGCTCTGCGGCGCTG 625
 Db 427 TGAAGGAGCGTGGGGCTACAGCTCTCTCTGAGAGACCGGAGCTCTGCGGCGCTG 486
 QY 626 AGCCCTCGGCGAGCTCTTGTGTGAGACCTGAGCCGCTGCGAGCTCATCTGTGTCTTT 685
 Db 487 AGCCCTCGGCGAGCTCTTGTGTGAGACCTGAGCCGCTGCGAGCTCATCTGTGTCTTT 546
 QY 686 CGAGAGCGCTTCTGAGAGCGGCTGTGTGAGACCTGCGGAGCGGCTGTGCTCGG 745
 Db 547 CGAGAGCGCTTCTGAGAGCGGCTGTGTGAGACCTGCGGAGCGGCTGTGCTCGG 606
 QY 746 TGTGAGAGCTACCCGAGAGCCATCTCATACCTTCAGAGGCGCAGAGCGGAGCGG 805
 Db 607 TGTGAGAGCTACCCGAGAGCCATCTCATACCTTCAGAGGCGCAGAGCGGAGCGG 666
 QY 806 CGCAGCCGCGCTCTCGCTCTGCGCAGACCGGAGCTGTGACCTTGTGTCTGSA 865
 Db 667 CGCAGCCGCGCTCTCGCTCTGCGCAGACCGGAGCTGTGACCTTGTGTCTGSA 726
 QY 866 GGGCGGCTCTGAGACTCTCTCTCGATTTTGAAGAGTGCAGTGGCTGCGCG 925
 Db 727 GGGCGGCTCTGAGACTCTCTCTCGATTTTGAAGAGTGCAGTGGCTGCGCG 786
 QY 926 GGAAGTGCAGTACAGGCGGTGGAAGAGACCCGAGAGCTGCAGAGCTGCAGAG 985
 Db 787 GGAAGTGCAGTACAGGCGGTGGAAGAGACCCGAGAGCTGCAGAGCTGCAGAG 846
 QY 986 ACCCATGTCATCTTCTGAGAGCGGAGTCCCTGAGAGCGGCGCTGAGCTGAGAGTGG 1045
 Db 847 ACCCATGTCATCTTCTGAGAGCGGAGTCCCTGAGAGCGGCGCTGAGAGTGG 906
 QY 1046 ACCGAGACCTGAGAGCGGAGTCTGAGGCGCTCTTTTGAAGAGCCATGACTC 1105
 Db 907 ACCGAGACCTGAGAGCGGAGTCTGAGGCGCTCTTTTGAAGAGCCATGACTC 966
 QY 1106 CACCGACACCAAGTGGGCTCTGCTGGGAGAGACCCGAGAGCTGCAGAGTGCAGTCTGG 1165
 Db 967 CACCGACACCAAGTGGGCTCTGCTGGGAGAGACCCGAGAGCTGCAGAGTGCAGTCTGG 1026
 QY 1166 ATTCGAGCTGCGAATACTACAGTCCGAGACACTTACAGCTGCTGGTGCAGAGATG 1225
 Db 1227 ATTCGAGCTGCGAATACTACAGTCCGAGACACTTACAGCTGCTGGTGCAGAGATG 1086
 QY 1226 ATATGTAG 1233
 Db 1087 ATATGTAG 1094

RESULT 7
 AAH9795/c

ID AAH9795 standard; cDNA: 851 BP.
 AC AAH9795;
 XX 16-OCT-2001 (first entry)
 DE Human protein encoding cDNA sequence spq ID NO:630.
 XX Human: cancer; HIV infection; human immunodeficiency virus;
 KW anti-infective; antipneumatic; antilethargic; immunosuppressive;
 KW anti-bacterial; endocrine; cardiac; central nervous system; vitruide;
 KW anti-HIV; fungicide; antimutagen; cardiovascular; antineurotic; anemic;
 KW antiaggregant; haemostatic; valnerary; antileuc; osteopathic; eczema;
 KW dermatological; antiallergic; antisthmatic; antiparkinsonian; infection;
 KW neuroprotective; antidepressant; noctropic; antiparkinsonian; inflammation;
 KW immunostimulant; gene therapy; antitense therapy; vaccine; pancreatitis;
 KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmune;
 KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
 KW chromocytopenia; osteoporosis; severe combined immunodeficiency;
 KW allergic rhinitis; diabetes; multiple sclerosis; depression;
 KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
 KW neurological disorder; ss.
 XX Homo sapiens.
 OS WO200153455-A2.
 XX 26-JUL-2001.
 XX 22-DEC-2000; 2000WO-US35017.
 PF 23-DEC-1999; 99US-0471275.
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Drmanac RT;
 PI WPI: 2001-457603/49.
 DR P-PSDB: AAM25854.
 XX Isolated human polynucleotides encoding polypeptides, useful for the
 XX treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
 XX Claim 1: Page 653; 1217p; English.
 XX AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
 CC AAM25963. The proteins can have activities based on the tissues and
 CC cells they are expressed in, such as: antineurotic; antilethargic;
 CC antilethargic; immunosuppressive; antibacterial; endocrine; cardiac;
 CC central nervous system; vitruide; anti-HIV; fungicide; antimutagen;
 CC cardiovascular; antineurotic; antisthmatic; antiparkinsonian; infection;
 CC antileuc; osteopathic; valnerary; antileuc; osteopathic; eczema;
 CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
 CC encoding them can be used in gene therapy, antitense therapy and vaccine
 CC production. The proteins and polynucleotides are useful for screening for
 CC agonists or antagonists of a protein and for the treatment and diagnosis
 CC of disorders associated with the activity of a protein e.g. inflammation,
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
 CC anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
 CC neurological disorders.
 CC Sequence 851 BP; 136 A; 273 C; 289 G; 153 T; 0 other;
 SO Query Match 41.1%; Score 506.4; DB 22; Length 851;

Best Local Similarity 99.8%; Pred. No. 5.9e-93;
Matches 507; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 726 CCGGGAGGGCCCTGCGCCGCTGTGAGACTCAACCCGACACCCATCTTATACACTTCA 785
DB 720 CAGGGAGGGCCCTGCGCCGCTGTGAGACTCAACCCGACACCCATCTTATACACTTCA 661
QY 786 GGGCCAGAGCCGAGACCCCGCCGCTCCGCTGTGCGCCAGACCCGACCT 845
DB 660 GGGCCAGAGCCGAGACCCCGCCGCTCCGCTGTGCGCCAGACCCGACCT 601
QY 846 GGTGACCTTGTCTGTGAGAGCCCGCTCCGCTGTGCGCCAGACCCGACCT 905
DB 600 GGTGACCTTGTCTGTGAGAGCCCGCTCCGCTGTGCGCCAGACCCGACCT 541
QY 906 AGTGCAGCTGGGGCTGCGCCGAGAGTGGGTACAGCCCGGTGAGAGACCCGAC 965
DB 540 AGTGCAGCTGGGGCTGCGCCGAGAGTGGGTACAGCCCGGTGAGAGACCCGAC 481
QY 966 GCAGCTGACGAGACGACAGACCCCATGCTGATTTCTGAGCCGAGCTGAGGGCCG 1025
DB 480 GCAGCTGACGAGACGACAGACCCCATGCTGATTTCTGAGCCGAGCTGAGGGCCG 421
QY 1026 GGGCCCTGACCTAGAGAGTGGAGCCCGACCTGAGGGCAGCTGGGTGTCGGGGCCCTGT 1085
DB 420 GGGCCCTGACCTAGAGAGTGGAGCCCGACCTGAGGGCAGCTGGGTGTCGGGGCCCTGT 361
QY 1086 TTTTGGAGAGCCATCATGCTCCACCCGACACACAGTGGGTGTCGGGTGAGAGAGCCGGAG 1145
DB 360 TTTTGGAGAGCCATCATGCTCCACCCGACACACAGTGGGTGTCGGGTGAGAGAGCCGGAG 301
QY 1146 CAGCGAAGTGAAGTCTCGATCTCGGCTCGGAAACTAGTGGCCGACAGACTTCTA 1205
DB 300 CAGCGAAGTGAAGTCTCGATCTCGGCTCGGAAACTAGTGGCCGACAGACTTCTA 241
QY 1206 CTGCTGTGTGTCAGAGATGATATGTAG 1233
DB 240 CTGCTGTGTGTCAGAGATGATATGTAG 213

RESULT 8
AAC06025
ID AAC06025 standard; cDNA; 249 BP.
XX
AC AAC06025;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 10100.
XX
KM Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KM gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP103401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 990US-0122487.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
```

PS Claim 1; SEQ ID 10100; 71pp + CD-ROM; English.

CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.

SQ Sequence 249 BP; 43 A; 87 C; 65 G; 54 T; 0 other;

Query Match 19.9%; Score 245.4; DB 21; Length 249;

Best Local Similarity 99.6%; Pred. No. 1.6e-40;
Matches 246; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 284 CCTCACCTGCTCATCCAGAACATCAGCTTCTCCTTCACTCTTACAGAGCTGGCC 343
DB 1 CCTCACCTGCTCATCCAGAACATCAGCTTCTCCTTCACTCTTACAGAGCTGGCC 60
QY 344 CTACAAAGCCAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 403
DB 61 CTACAAAGCCAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 120
QY 404 CCGCCCTGCTATGTCAGAGTCCCGTCTCAGCTGCTGTGTGACAGAGCGGTATG 463
DB 121 CCGCCCTGCTATGTCAGAGTCCCGTCTCAGCTGCTGTGTGACAGAGCGGTATG 180
QY 464 GGGAGGTGAGATTAACGACGGGAAGCTCTACAGCCCTACCTCTTACAGCGACTGCC 523
DB 181 GGGAGGTGAGATTAACGACGGGAAGCTCTACAGCCCTACCTCTTACAGCGACTGCC 240
QY 524 CCGAGGA 530
DB 241 CCGAGGA 247
```

RESULT 9

AAK61892
ID AAK61892 standard; cDNA; 266 BP.

XX
AC AAK61892;

XX
DT 06-NOV-2001 (first entry)

XX
DE Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:6952.

XX
KM Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KM cytostatic; gene therapy; vaccine; metastasis; ss.

XX
OS Homo sapiens.

XX
PN WO200157182-A2.

XX
PD 09-AUG-2001.

XX
PF 17-JAN-2001; 2001WO-US01354.

XX
PR 31-JAN-2000; 2000US-0179065.

XX
PR 04-FEB-2000; 2000US-0180628.

XX
PR 24-FEB-2000; 2000US-0184664.

XX
PR 02-MAR-2000; 2000US-0186350.

XX
PR 16-MAR-2000; 2000US-0189874.

XX
PR 17-MAR-2000; 2000US-0190076.

XX
PR 18-APR-2000; 2000US-0198123.

XX
PR 19-MAY-2000; 2000US-0205515.

Dd		732	CCTCCTGTCGTCCCGGCCCTTGCATGCTGGCCCTCCTGCTCCTAGCCCCCTCCTGGCCCTTC	673
Oy		733	GACGCCCTCATGCTGCTTTTGCGAGACCCTTCTGTGAACCGGGCGGTGAGCAAGCAACT	724
Dd		672	TGCACCTCCTCTCGTGCCTGCCCTCCTCCTCTCCTGCACCCCTCGGCCCTTCCTGCCCCTTC	613
Oy		725	TCCGGGAGGGGCGCTGTGCCGGCTGCTGAGACTCACCCGACAACCATTTCATACCTTCG	784
Dd		612	TGCTCCTGGCCCCCTCCTCGCTGCTGCTGCCCCCTCTCGCCCCCTCGGCCCTCCTCCTGCTGC	553
Oy		785	AAGGCACAGAGGGCGACCCCGCGACCCGGCGCTGCGCTGCGCGACACACCGCCACG	844
Dd		552	TGCCCTCTCCTGCCCCCTCTCTGCTCTGCTCTGCTGCCCCCTCTCCTGCTCTGCCCCCTTCG	493
Oy		845	TGCTGACCTTGCTGCTGCTGAGAGCCC	870
Dd		492	TCCTGCCCCCTCCTCGCTCGCTGCCCC	467
RN		13		
AAFB		2902/C		
ID		AAF82902	standard; DNA; 1926 BP.	
XX		AAF82902;		
DE		29-JUN-2001	(first entry)	
XX				
DE			EBV tethering protein EBNA1 encoding DNA.	
KW			Histone H1; tethering protein; LANA; gene therapy; multiple sclerosis;	
KM			Parkinson's disease; Huntington disease; diabetes; human herpesvirus 8;	
KW			EBV latency-associated nuclear antigen; LANA; EBNA1; ds.	
XX				
OS			Epsstein-Barr virus.	
XX				
FN				
FT				
Key			Location/Qualifiers	
CDS			1..1926	
/tag=		a		
MO200125484-A2.				
PD			12-APR-2001.	
PF			29-SEP-2000; 2000WO-US26908.	
PR			01-OCT-1999; 99US-0410399.	
JNM1) UNITV MICHIGAN.				
Robertson ES, Colter MA;				
WIPI: 2001-281736/29.				
P-PSTD; AAB62332.				
A composition for use in gene therapy comprises an expression vector				
that includes a nucleic acid sequence encoding a nucleic acid binding				
protein -				
Disclosure: Fig 9C; 60pp. English.				
The invention provides a composition comprising nucleic acid, histone H1				
protein and expression vector operationally encoding a protein suitable				
for tethering the nucleic acid to the histone H1 protein, where the				
tethering protein is LANA. The composition is useful in aiding the				
retention of the viral DNA in the host cell. The viral vector encodes a				
protein suitable for tethering DNA to Histone H1. Methods for screening				
for compounds which are agonistic or antagonistic for the tethering of				
viral proteins to histone H1 and DNA binding sites are useful for				
developing the method of viral transfer. The composition has applications				
to gene therapy, including the treatment of multiple sclerosis,				
Parkinson's disease, Huntington disease and diabetes. The present				
sequence represents the nucleotide sequence of the Epstein-Barr virus				
(EBV) tethering protein ENNA1.				

XX	Sequence	1926 BP; 487 A; 352 C; 872 G; 215 T; 0 other:
Query Match	4.6%; Score 56.4; DB 22; Length 1926;	
Best Local Similarity	44.5%; Pred. No. 0.021;	
Matches 225; Conservative	0; Mismatches 281; Indels	0; Gaps
OY	365 TCCTGGCCCTCCTCGTGTGTCGTGGGCCCTGCTGTGGCGCCCTGCTATATGCAGT	424
DB		913
OY	972 TCTCTCGGCCCTCCMGCTCGCCCCCTTCCTGTCTCTGCCCCTCGGCCCTCGCC	913
DB	425 GGCGTCAAGCTGCTGCTGTGTACGAGAGCGCATAGGAGAGGTGAAGAACAAGC	484
DB	912 TCTTGCCCCCTCGGCCCTCGTCTGTGCCCCCTCTGCCCCCTGCTGTCTGCCCC	853
OY	485 GGAAGCTACACACGCTACGCTCTACTACAGAGATGCCCCGAGAGCGAAGTTCGTA	544
DB	852 TGCCCCCTCTCTGTGCTCTGCCCCCTCTGCCCCCTCTCTCTGCTCTGCCCCCT	793
OY	545 ACTTCATCTTAAAGCGCAGCTGAGAGCGCGCTGCGGAGCTAACAGCTTTCTGAGAC	604
DB	792 TCTGTGCTGTGCCCCCTCGGCCCTCTGCTGTGCCCCCTCTGCCCCCTCGCTCTG	733
OY	605 GGGACCTCTGCGCGGCGCTGAGCCCTCGCGGACCTCTTGTGTAACCTGAGCCGCTGC	664
DB	732 CCTCTGCTGTGCCCCCTCGGCCCTCTGCTGTGCCCCCTCTGCTGTGCCCCCT	673
OY	665 GACGCTCTATCGTGTGCTGTGCGAGCGCTTCTGAGCGGCGCGCTGATGACCCACAGT	724
DB	672 TGCCCCCTCTCTGTGCTCTGCCCCCTCTGTCTGTCTGTGCCCCCTCTGCCCCCT	613
OY	725 TCCGGAGGAGCGCTGTGCGCGCTGTGAGAGTCACCOCGAGACCATTCTATCACCTTCG	784
DB	612 TCTCTGTGCCCCCTCGGCCCTCTGCTGTGCCCCCTCTGCCCCCTCTGCTGTG	553
OY	785 AAGGCGCAGAGGCGCGACCCCGGCACCGGCGCTCGCTGTGCGCAGACCGGCACG	844
DB	552 TGCCCCCTCTGCGCCCCCTCGCTGTGCGCCCCCTCTGCTGTGCCCCCTCTGCG	493
OY	845 TGGTAGCTTGTGCTGTGAGAGGCC 870	
DB	492 TCTGTGCCCCCTCTGTGCTGTGCCCC 467	
RESULT 14		
AAA75454/C		
ID AAA75454 standard; DNA; 2580 BP.		
XX AAA75454;		
DT 15-JAN-2001 (first entry)		
DE Nucleotide sequence of the Epstein Barr nuclear antigen.		
XX Origin of replication; protein-protein interaction; replication;		
KW two-hybrid system; nuclear antigen; ss.		
OS Epstein-barr virus.		
XX US6114111-A.		
PN 05-SEP-2000.		
PD 30-MAR-1998; 98US-0050863.		
PF 30-MAR-1998; 98US-0050863.		
PR (RIGE-) RIGEL PHARM INC.		
PA Luo Y, Payan D, Huang B;		
XX WPI; 2000-593546/56.		
DR		

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2002, 18:26:01 : Search time 62 Seconds
(without alignments)
734.520 Million cell updates/sec

Title: US-09-598-443-2

Perfect score: 2147

Sequence: 1 MPGVCDRAPFLSPEDQVL.....GSRNYSARTDFYCLVSKDM 410

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Sequences: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: A_Geneseq_032802.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2147	100.0	410	AAV25426	Human SIGIRR prote
2	2147	100.0	410	AAAB61139	Human NOV9 protein
3	1761	82.0	407	AAU17408	Novel signal trans
4	1761	82.0	407	AAU43652	Human polypeptide
5	1761	82.0	410	AAU43583	Human polypeptide
6	1749	81.5	363	AAAB58416	Lung cancer associ
7	974.5	45.4	212	AAU25854	Human protein sequ
8	267.5	12.5	658	AAU22164	Human TIGIRK prote
9	267.5	12.5	686	AAU4128	Human IL-1R2 prot
10	257.5	12.0	579	AAU14130	Human IL-1R2 prot
11	256.5	11.9	481	ABBI0347	Human CDNA SEQ ID

12	256.5	11.9	486	AAU14129	Human IL-1R210 pro
13	256.5	11.9	696	AAU95298	Human Xrec2, Homo
14	256.5	11.9	696	AAU78808	Human protein SEQ
15	256.5	11.9	710	ABBI1785	Human oligophrenin
16	256.5	11.9	710	AAU79792	Human protein SEQ
17	254	11.8	519	AAU77140	Murine interleukin
18	254	11.8	519	AAU39781	Interleukin-18 rec
19	254	11.8	537	AAU31908	Mouse receptor pro
20	254	11.8	537	AAU6615	Mouse interleukin-
21	232	10.8	521	AAU39780	Interleukin-18 rec
22	232	10.8	540	AAU77145	Human interleukin-
23	232	10.8	540	AAU39786	Interleukin-18 rec
24	231.5	10.8	541	AAU31907	Human receptor pro
25	231.5	10.8	541	AAU6613	Human interleukin-
26	229.5	10.7	570	AAU01911	Human interleukin-
27	215.5	10.0	570	AAU01913	Mouse interleukin-
28	209.5	9.8	553	AAU42060	IL-1R, Homo sapie
29	209.5	9.8	569	AAU92001	Derived sequence o
30	209.5	9.8	569	AAU90330	Human interleukin-
31	209.5	9.8	569	AAU59090	Human interleukin-
32	209.5	9.8	569	AAU60870	Human IL-1 recepto
33	209.5	9.8	569	AAU30376	Human interleukin-
34	209.5	9.8	569	AAU22733	Human interleukin-
35	209.5	9.8	569	AAU06541	Human interleukin-
36	209.5	9.8	569	AAU37791	Human interleukin-
37	209.5	9.8	569	AAU30921	A human interleuk
38	209.5	9.8	569	AAU59700	Human interleukin-
39	204.5	9.5	561	AAU31065	Rat interleukin-1
40	204.5	9.5	561	AAU06665	Mouse interleukin-
41	204.5	9.5	561	AAU03539	Rat interleukin-1
42	204	9.5	562	AAU31064	Human interleukin-
43	204	9.5	562	AAU06664	Human interleukin-
44	204	9.5	562	AAU03538	Human interleukin-
45	202	9.4	556	AAU00275	Transmembrane form

ALIGNMENTS

RESULT 1	AAV25426	standard; Protein: 410 AA.
XX	AAV25426;	
XX	AC	
XX	08-SEP-1999	(first entry)
XX	DE	Human SIGIRR protein.
XX	KW	SIGIRR; anti-inflammatory; anti-autoimmune disease; chromosome 11;
KW	interleukin-1 receptor; 11p15.5; arthrogryposis multiplex congenita;	
KW	breast cancer; insulin-dependent diabetes; sickle cell anemia;	
KW	bladder cancer; detection; cell-signal transduction; gene therapy;	
KW	immune regulation; cell proliferation; cell death; cell migration;	
KW	cell interaction; inflammation; marker; mass spectrometry; human.	
XX	OS	Homo sapiens.
XX	PN	W09932626-A1.
XX	PD	01-JUL-1999.
XX	PE	23-DEC-1998; 98WO-US27368.
XX	PR	23-DEC-1997; 97US-0068770.
XX	PA	(IMMUNEX CORP.
XX	PI	Sims JE; /
XX	WT	WPI; 1999-418925/35.
DR	N-PSDB;	AAU88091.
XX		

PT New interleukin-1 receptor analog SIGIRR nucleic acid and proteins
XX
XX Claim 2; Page 69-71; 72pp; English.

CC This invention describes a novel human SIGIRR DNA, its allelic variants
CC or species homologues which have anti-inflammatory and anti-autimmune
CC disease activity. SIGIRR is an analog of interleukin-1 receptor. The
CC products of the invention are used (a) as probes or primers for
CC identifying human chromosome 11, to map genes on this chromosome and to
CC identify disease-related genes (particularly in the region 11p15.5 where
CC genes are present associated with e.g. arthrogryposis multiplex
CC congenita, breast cancer, insulin-dependent diabetes, sickle cell
CC anaemia, bladder cancer), including detection of defective genes; (c)
CC to study cell-signal transduction and the SIGIRR system, and (d) in
CC gene therapy. Sense and antisense oligonucleotides can be used to inhibit
CC expression of the SIGIRR gene. The proteins of the invention are used:
CC (i) to study cellular processes (immune regulation, proliferation, death,
CC migration, interaction with other cells and inflammation); (ii) to
CC identify and purify proteins that associate with SIGIRR ligands and
CC receptors, and to measure their biological activity; (iii) in screening
CC for, and rational design of, potential inhibitors of activity; (iv)
CC therapeutically against diseases mediated by SIGIRR polypeptide
CC counter structures; (v) as molecular weight (m.w.) markers in
CC electrophoresis; (vi) for determining isoelectric points of unknown
CC proteins; (vii) as controls for determining the extent of protein
CC fragmentation (e.g. to aid characterization of protein structures by
CC mass spectrometry); (viii) for generation of antibodies (Ab); (ix) to
CC deliver diagnostic or therapeutic agents to cells that express SIGIRR
CC binding molecules.

XX
XX SQ: Sequence 410 AA:

Query Match 100.0%; Score 2147; DB 20; Length 410;
Best Local Similarity 100.0%; Pred. No. 7.6e-220;
Matches 410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPGVCDRAPDFLSPEDQVLRPALGSSVALNCTAMVYSGPHCSIPSYQMKDGLPGIGG 60
DB 1 MPGVCDRAPDFLSPEDQVLRPALGSSVALNCTAMVYSGPHCSIPSYQMKDGLPGIGG 60
QY 61 HYSLHEYSWKANLSEVLVSSVGVNTSTEVYGAFTCSIONISFSSFTLQRAGPTSHVA 120
DB 61 HYSLHEYSWKANLSEVLVSSVGVNTSTEVYGAFTCSIONISFSSFTLQRAGPTSHVA 120
QY 121 AVLASLVLALLAALLVYKCRNLVLMYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180
DB 121 AVLASLVLALLAALLVYKCRNLVLMYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180
QY 181 VNFILKPOLERRRGYKFLDQDRLPRAEPSADLVNLSRCRRLIVLSDAFLSRAMCSH 240
DB 181 VNFILKPOLERRRGYKFLDQDRLPRAEPSADLVNLSRCRRLIVLSDAFLSRAMCSH 240
QY 241 SFREGICRLLELRRPLFTFEFGORDPAHPALRLRQHRHLVTLMLRPGSVTPSSDFW 300
DB 241 SFREGICRLLELRRPLFTFEFGORDPAHPALRLRQHRHLVTLMLRPGSVTPSSDFW 300
QY 301 KEVQOLAPRKVRYRPVPGDQDQLODDKDPMLILGRVPEGALDSEVDPPEGLGVRG 360
DB 301 KEVQOLAPRKVRYRPVPGDQDQLODDKDPMLILGRVPEGALDSEVDPPEGLGVRG 360
QY 361 PVGGEPAHPHTSGVSLGESRSSEVDVSDLSGRNRSARDFECLVSKDDM 410
DB 361 PVGGEPAHPHTSGVSLGESRSSEVDVSDLSGRNRSARDFECLVSKDDM 410

RESULT 2
AAB61139
ID AAB61139 standard; Protein: 410 AA.
XX
AC AAB61139;
XX
DT 30-MAR-2001 (first entry)

XX
DE Human NOV9 protein.
XX
KM Human; NOVX; antiinflammatory; cytosstatic; neuroprotective;
KM cerebroprotective; immunomodulator; vunerary; vasotropic; gene therapy;
KM hyperplasia; tumor; resenosis; psoriasis; Dupuytren's contracture;
KM diabetes; Rheumatoid arthritis; cerebral oedema; Alzheimer's disease.
XX
OS Homo sapiens.
XX
PN WO200075321-A2.
XX
PD 14-DEC-2000.
XX
PF 01-JUN-2000; 2000WO-US15303.
XX
PR 03-JUN-1999; 99US-0137322.
PR 16-MAR-2000; 2000US-0189810.
PR 22-MAR-2000; 2000US-0191158.
PR 30-MAR-2000; 2000US-0193086.
PR 31-MAY-2000; 2000US-0137322.
XX
PA (CURA-) CUBAGEN CORP.
XX
PI Shinkets RA, Fernandes E, Herrman J, Vernet C;
XX
XX WPI: 2001-102403/11.
DR N-PSDB; AAF27857.
XX
PT New NOVX polypeptides and polynucleotides, useful in gene therapy, as a
PT diagnostic marker, protein therapeutic, antibody or small molecule drug
PT target for treating immune, proliferative and metabolic diseases and
PT wound healing -
XX
PS Claim 1; Page 36-38; 194pp; English.
XX
XX The present sequence is a new isolated polypeptide (NOVX). The NOVX
CC polypeptides, NOVX nucleic acids, and anti-NOVX antibodies are useful for
CC treating or preventing NOVX-associated disorders. They are also useful
CC for determining the presence of or a predisposition to a disease
CC associated with altered levels of the NOVX polypeptide or nucleic acid.
CC These NOVX-associated disorders include hyperplasias, tumors,
CC resenosis, psoriasis, Dupuytren's contracture, diabetic complications,
CC Rheumatoid arthritis, cerebral lesions, diabetic neuropathies, cerebral
CC oedema, senile dementia or Alzheimer's disease. The NOVX polynucleotides
CC are especially useful in gene therapy. Specifically, NOVX is useful as
CC a diagnostic marker or prognostic marker, protein therapeutic and
CC antibody target or small molecule drug target to treat disorders in the
CC immune response pathway, thyroid and metabolic diseases, bone metabolic
CC disorders, diseases of the pancreas (e.g. diabetes or digestive
CC disorders), proliferative diseases, or tissue regeneration and
CC development (e.g. wound healing or treatment of burns).
XX
SQ Sequence 410 AA:
Query Match 100.0%; Score 2147; DB 22; Length 410;
Best Local Similarity 100.0%; Pred. No. 7.6e-220;
Matches 410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPGVCDRAPDFLSPEDQVLRPALGSSVALNCTAMVYSGPHCSIPSYQMKDGLPGIGG 60
DB 1 MPGVCDRAPDFLSPEDQVLRPALGSSVALNCTAMVYSGPHCSIPSYQMKDGLPGIGG 60
QY 61 HYSLHEYSWKANLSEVLVSSVGVNTSTEVYGAFTCSIONISFSSFTLQRAGPTSHVA 120
DB 61 HYSLHEYSWKANLSEVLVSSVGVNTSTEVYGAFTCSIONISFSSFTLQRAGPTSHVA 120
QY 121 AVLASLVLALLAALLVYKCRNLVLMYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180
DB 121 AVLASLVLALLAALLVYKCRNLVLMYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180
QY 181 VNFILKPOLERRRGYKFLDQDRLPRAEPSADLVNLSRCRRLIVLSDAFLSRAMCSH 240
DB 181 VNFILKPOLERRRGYKFLDQDRLPRAEPSADLVNLSRCRRLIVLSDAFLSRAMCSH 240

Db 181 VNFILKPOLERRRGYKFLDLDRLLPRAEPSADLLVNI.SRCRLLIYVLSDAFLSRANC5H 240
QY 241 SFRGGLCRLELETRRPPIFTFEGORRDPAPRALRLRQHRHLVTLTMRPGSVTPSSDFW 300
Db 241 SFRGGLCRLELETRRPPIFTFEGORRDPAPRALRLRQHRHLVTLTMRPGSVTPSSDFW 300
QY 301 KEVQALPRKRVRRYPVGEDPOTQLQDDKDPMLILRGVRPGRALDSFVDPPEGLGVRG 360
Db 301 KEVQALPRKRVRRYPVGEDPOTQLQDDKDPMLILRGVRPGRALDSFVDPPEGLGVRG 360
QY 361 PVGEPSPAPHTSGVSLGESRSSEVDVSLGSRNYSARTDPYCLVSKDM 410
Db 361 PVGEPSPAPHTSGVSLGESRSSEVDVSLGSRNYSARTDPYCLVSKDM 410

RESULT 3
AAU17408
IT AAU17408 standard; Protein; 407 AA.
X AC AAU17408;
XX
DT 07-NOV-2001 (first entry)
XX
DE Novel signal transduction pathway protein, Seq ID 973.
XX
XX Neuroprotective; cytosolic; dermatological; immunosuppressive; tumour;
KM antinflammatory; anti-HIV; antibacterial; antinflammatory; cancer;
KM immune system disorder; rheumatoid arthritis; inflammatory condition;
KM organ transplant rejection; infection; hepatitis C; blood disorder;
KM sickle cell anemia; hyperproliferative disorder; Gaucher's disease;
KM neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KM chromosomal abnormality; Down syndrome; ischemia; renal disorder;
KM cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KM reproductive system; gastrointestinal; liver disorder; AIDS;
KM acquired immune deficiency syndrome.
XX
XX Homo sapiens.
OS
XX
PN MO200154733-A1.
XX
PD 02-AUG-2001.
XX
XX
PF 17-JAN-2001; 2001MO-US01312.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0227009.
PR 01-SEP-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234597.
PR 25-SEP-2000; 2000US-0234598.
PR 25-SEP-2000; 2000US-0234599.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.

PS Claim 1, SEQ ID No. 973: 860pp: English.

XX

CC The invention relates to novel isolated polypeptides (I), and
CC polynucleotides (II), (I), (II) and the antibody to (I) are useful for
CC diagnosing, preventing and treating diseases including immune system
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
CC transplant rejections and graft versus host disease, infectious diseases
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
CC other blood-related disorders (sickle cell anaemia), myeloproliferative
CC disorders, primary haematopoietic disorders, hyperproliferative
CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
CC disorders (e.g. Alzheimer's disease and Parkinson's disease), chromosomal
CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
CC disorders (e.g. glomerulonephritis), cardiovascular disorders
CC (e.g. arrhythmia), respiratory disorders, dermatological disorders
CC (e.g. healing, epithelial cell proliferation, endocrine disorders (e.g.
CC Addison's disease), reproductive system disorders, gastrointestinal
CC disorder (inflammatory disorders), liver disorders (cirrhosis),
CC as stimulators of B-cell responsiveness to pathogens, activators of
CC T-cells, to induce higher affinity antibodies, and as a means to induce
CC tumour proliferation in pathologies e.g. acquired immune deficiency
CC syndrome (AIDS). AAU17059-AAU17683 represent novel signal transduction
CC pathway protein, amino acid sequences of the invention.

	Best Local Similarity	100.0%:	Pred. No. 1,1e-178:	Mismatches	0:	Indels	0:	Gaps	0:
	Matches	341:	Conservative	0:	Mismatches	0:	Indels	0:	Gaps
OY	70	VKANLSEVLYSVVLGNVNTSTFVGAPTCISIQNIISFSFTLQRAQPTSHVAALASLVLL	129						
Dd	67	VKRALSEVLSSVSLGNNVTSTFYAFCTSIQNIISFSFTLQRAQPTSHVAALASLVLL	126						
OY	130	LALLLAALLYVKRCLNVLLMWODAYEVEINDGKLVDAYYSVSDCPEDKREVNFIILKPOL	189						
Dd	127	LALLLAALLYVKRCLNVLLMWODAYEVEINDGKLVDAYYSVSDCPEDKREVNFIILKPOL	186						
OY	190	ERRRGKYKLFDDRDILLPRAEPSADLIIVLNSCRRLIYVLSDAFLSRAMCSHSPREGICRL	249						
Dd	187	ERRRGKYKLFDDRDILLPRAEPSADLIIVLNSCRRLIYVLSDAFLSRAMCSHSPREGICRL	246						
OY	250	LELTRRRIFFTFEGRORDPANHALLRLRLQHHLVTLMLLPBGSAVTPSSDPDWKEVOLATLR	309						
Dd	247	LELTRRRIFFTFEGRORDPANHALLRLRLQHHLVTLMLLPBGSAVTPSSDPDWKEVOLATLR	306						
OY	310	KVRPRPVGGPOQTOLDOKDPMILIRGVNPGRALDSVPDPDEGDIGVGPVFGEPSAP	369						
Dd	307	KVRPRPVGGPOQTOLDOKDPMILIRGVNPGRALDSVPDPDEGDIGVGPVFGEPSAP	366						
OY	370	PHTSGSVIGESRSSEVDVSDLSRNYSARTPFYCYSKDDM	410						
Dd	367	PHTSGSVIGESRSSEVDVSDLSRNYSARTPFYCYSKDDM	407						
	RESULT: 4								
	AAM43652								
ID	AAM43652 standard; Protein; 407 AA.								
XX	AAM43652;								
XX									
DT	22-OCT-2001 (first entry)								
XX									
DE	Human polypeptide SEQ ID NO 330.								
KW	Human; antiarthritic; antirheumatic; antiproliferative; vasotropic;								
KW	cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;								
KW	fungicide; ophthalmological; cytostatic; immunosuppressive; nootropic;								
KW	neuroprotective; antiallergic; hepatotropic; antidyspeptic;								
KW	antiinflammatory; antiulcer; vulnerary; anticonvulsant; antibacteri-								
KW	antiparasitic; cardiant; gene therapy; cancer; immune disorder;								
KW	cardiovascular disorder; neurological disease; infection; human.								
OS	Homo sapiens.								
PN	WO200155308-A2.								
PD	02-AUG-2001.								
XX									
PF	17-JAN-2001: 2001MO-US01309.								
XX									
PR	31-JAN-2000: 2000US-0179065.								
PR	04-FEB-2000: 2000US-0180628.								
PR	24-FEB-2000: 2000US-0184664.								
PR	02-MAR-2000: 2000US-0186350.								
PR	16-MAR-2000: 2000US-0188874.								
PR	18-MAR-2000: 2000US-0190076.								
PR	18-APR-2000: 2000US-0198123.								
PR	19-MAY-2000: 2000US-0205515.								
PR	07-JUN-2000: 2000US-0209467.								
PR	28-JUN-2000: 2000US-0214886.								
PR	30-JUN-2000: 2000US-0215135.								
PR	07-JUL-2000: 2000US-0216880.								
PR	11-JUL-2000: 2000US-0217487.								
PR	11-JUL-2000: 2000US-0217496.								
PR	14-JUL-2000: 2000US-0218290.								
PR	26-JUL-2000: 2000US-0220963.								
PR	26-JUL-2000: 2000US-0220964.								
PR	14-AUG-2000: 2000US-0224518.								

XX Lung cancer associated gene sequences, referred to as lung cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as lung cancer -
XX
XX
PS Claim 11: Page 1271-1272; 1425pp; English.
XX
XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
CC associated proteins represented in AA558106 - AAB58548 lung cancer
CC associated proteins and polynucleotide sequences, their agonists, and
CC antagonists may have neuroprotective, cytostatic, cardiocative, and
CC immunomodulatory, muscular active, general, gynecological, urological, gastrointestinal
CC general, nephrotropic, antiinfective, antimicrobial, or antibacterial
CC activity. The invention also includes antibodies specific for the
CC protein or polynucleotide sequences. The lung cancer associated
CC polynucleotide sequences may be used for detection of lung cancer.
CC Chromosome identification, as chromosome markers, and for numerous other
CC diagnostic or research purposes. The proteins may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders. The proteins may also be used in the treatment of wounds and
CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and
CC peptide AA558549 are used in the course of the invention for the
CC identification and characterisation of the polynucleotide and protein
CC sequences.
XQ Sequence 363 AA;

Query Match	81.58;	Score 1749;	DB 21;	Length 363;
Best Local Similarity	99.48;	Pred. No. 1.7e-177;		
Matches 339;	Conservative	0;	Mismatches 2;	Indels 0;
				Gaps 0;

Qy	70	VANISEVYSSVLCVANTSTVEY	GATFCSTIONISSESTFIORAGTSHVAVALSLIYL	12
Db	23	VKANISEVLYSVLCVANTSTVEY	GATFCSTIONISSESTFIORAGTSHVAVALSLIYL	82
Qy	130	LALLAALLYYVCRLNYLWQADAYGVEINODKIYDAVVSQDEPRKPYNFIKROL	18	
Db	83	LALLAALLYYVCRLNYLWQADAYGVEINODKIYDAVVSQDEPRKPYNFIKROL	142	
Qy	190	ERRRGKFLDRODRLPAEPESADLYNISRCRLIYLSYDAFLSAMCSHFREGICRL	24	
Db	143	ERRRGKFLDRODRLPAEPESADLYNISRCRLIYLSYDAFLSAMCSHFREGICRL	202	
Qy	250	LELTRRPFITTEGCRDPAPARLRLKORHLVTLTLMPGCVTPSSDFMKENYOLALPR	309	
Db	203	LELTRRPFITTEGCRDPAPARLRLKORHLVTLTLMPGCVTPSSDFMKENYOLALPR	2623	
Qy	310	KARYRVEEDPOTQLODDKDPMLILKRGVPEGRALDEVPDEPDGLVGRGVTEGPSAP	365	
Db	263	KARYRVEEDPOTQLODDKDPMLILKRGVPEGRALDEVPDEPDGLVGRGVTEGPSAP	3223	
Qy	370	PHRSGVSLGESRSSSEVNDVSLGSRBNYSARTIDFYCLVSKDDM	410	
Db	323	PHRSGVSLGESRSSSEVNDVSLGSRBNYSARTIDFYCLVSKDDM	363	

DE	XX	AC	XX	DT	XX	SE
Human protein sequence	AA025854	AA025854	16-OCT-2001	(first entry)	SE0	ID NO:1369

KM neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
 KM immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
 KM antiapoptotic; rheumatoid arthritis; septic shock; pancreatitis;
 KM cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmune;
 KM genetic disease; haematopoietic disorder; platelet disorder; asthma;
 KM thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
 KM allergic rhinitis; diabetes; multiple sclerosis; depression;
 KM Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
 KM neurological disorder.

PS Claim 20: Page 283; 1217P; English.

XX

CC AAM99166 to AAM99904 encode the human proteins given in AAM25225 to

CC AAM25963. The proteins can have activities based on the tissues and

CC cells they are expressed in, such as: antinflammatory; antipneumatic;

CC antithalctic; immunosuppressive; antibacterial; endocrine; cardiac;

CC central nervous system; vincide; anti-HIV; fungicide; antimutagen;

CC cardiovascular; antinaemic; antigravty; haemostatic; vulnerary;

CC antitumor; osteoplastic; neuroprotective; antiallergic; antistaphic;

CC antidiabetic; cytostatic; demutagenic; antidiabetic; nootropic;

CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides

CC encoding them can be used in gene therapy, antisense therapy and vaccine

CC production. The proteins and polynucleotides are useful for screening for

CC agonists or antagonists of a protein and for the treatment and diagnosis

CC of disorders associated with the activity of a protein e.g. inflammation,

CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,

CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal

CC infections, autoimmunity, genetic diseases, haematopoietic disorders,

CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,

CC osteoporosis, severe combined immunodeficiency, eczema, allergic

CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,

CC Alzheimer's disease, Parkinson's disease, neurodegenerative and

CC neurological disorders.

XX

Sequence 212 AA:

Query Match	45.48;	Score 974.5;	DB 22;	Length 212;
Best Local Similarity	89.28;	Pred. No. 3e-95;		
Matches 189;	Conservative	0;	Mismatches	0;
			Indels	23;
			Gaps	1

QY 379 ESRSEVDVSDGSRNYSARTDFYCLVSKDDM 410
 DB 181 ESRSEVDVSDGSRNYSARTDFYCLVSKDDM 212

RESULT 8
 ID AAY22164
 AAY22164 standard; Protein; 658 AA.

AC AAY22164;

DT 08-SEP-1999 (first entry)

DE Human TIGIR protein sequence.

XX TIGIR; human; chromosome X; FG syndrome; premature ovarian failure-1;
 KM telomomatosi; epilepsy; Bazex syndrome; detection; defective gene;
 KM cell-signal transduction; gene therapy; inhibitor; immune regulation;
 KM cell proliferation; inflammation.

XX Homo sapiens.

PN WO932629-A1.

PD 01-JUL-1999.

PF 23-DEC-1998; 98WO-US27625.

PR 23-DEC-1997; 97US-0068634.

PA- (IMMV) IMMUNEX CORP.

XX Sims JE;

XX WPI; 1999-418928/35.

DR- N-PSDB; AAX84308.

PT New interleukin-1 receptor analog TIGIR nucleic acid and proteins
 PT used to, e.g. treat autoimmune disease

PS Claim 2; Page 10; 79pp; English.

XX This sequence is the human TIGIR protein of the invention. The
 CC TIGIR DNA sequences can be used: (a) as probes or primers for
 CC identifying nucleic acid that encodes proteins with TIGIR activity;
 CC (b) to identify human chromosome X; to map genes on this chromosome
 CC and to identify disease-related genes (particularly in the region
 CC Xq21.3-22 where genes are present associated with e.g. FG syndrome,
 CC premature ovarian failure-1, telomomatosi, epilepsy, Bazex syndrome
 CC etc.) including detection of defective genes; (c) to study cell-signal
 CC transduction and the TIGIR system; and (d) in gene therapy. Sense and
 CC antisense oligonucleotides derived from the TIGIR coding sequence can be
 CC used to inhibit expression of the TIGIR gene. The TIGIR protein, or its
 CC soluble fragments, are used: (i) to study cellular processes (immune
 CC regulation, proliferation, death, migration, interaction with other cells
 CC and inflammation); (ii) to identify and purify proteins that associate
 CC with TIGIR ligands and receptors, and to measure their biological
 CC activity; (iii) in screening for, and rational design of, potential
 CC inhibitors of activity; (iv) therapeutically against diseases mediated by
 CC TIGIR polypeptide counter-structures; (v) as molecular weight markers in
 CC electrophoresis; (vi) for determining isoelectric points of unknown
 CC proteins; (vii) as controls for determining the extent of protein
 CC fragmentation (e.g. to aid characterisation of protein structures by mass
 CC spectrometry); (viii) for generation of antibodies (Ab); and (ix) to
 CC deliver diagnostic or therapeutic agents to cells that express TIGIR
 CC binding molecules. Ab are used for affinity purification of TIGIR;
 CC therapeutically to inhibit binding of TIGIR to its counter structures,
 CC and (if agonistic) to promote cell signalling.

XX Sequence 658 AA;

Query Match 12.5%; Score 267.5; DB 20; Length 658;

Best Local Similarity 27.2%; Pred. No. 4,2e-19;
 Matches 103; Conservative 72; Mismatches 143; Indels 60; Gaps 19;

QY 6 DRAPDPLSSEDD--VLRPALSSVALNCTAWY-VSGPHOSLPSYQWMLK-DGLPLIGGH 61
 DB 209 DKPPKPLFPEWNPVVIDVOLGKPLNIPCKAFPGFSGE--SGPMIYMWGGEKFIIELAGH 266
 QY 62 YSLHEPFWKANDSE-----VYSSVLGVNVTSTVEYGAFTCSIONIS--FSSFTLOR 112
 DB 267 IRGGEIRLKEHGEKEVEELALIFDSVENDLAN-----YTHVENRNKRKIASVLLRK 320
 QY 113 AGPTSHV--AAVLASLLVLLALLLVKCRNLNVLMYQDAYGEVEIN-DGKLYDAYV 169
 DB 321 KDLIYKIELAGLGALIFLLVLLV--VIYKCYNIEFLMFRQHFADENNDKREYDAYL 378
 QY 170 SYS-----DC--PEDRKFNFIILKPOLERRRGVLFDDRDRLRAESADLYNLNR 220
 DB 379 SYTKVDQDTLDCNPEEDQALEVLPDVELEKHGYKLFPERDLPSGYMEDLTGYVEDQ 438
 QY 221 CRLLIYVLS-DAFLSRAMCSHSFREGICRL--ELTRRPIFTFEGORRDPAPALRL 276
 DB 439 SRLLIIVLPDYTLIRKGSIFELSRNLHNLVSGEL--KYIILECTELKGVNCOQVESI 496
 QY 277 RQHRHLVTLILMRPGSVTP--SSDFKKEVQALPRKVRVRPVEGDPTOLQDDKDPMLIL 334
 DB 497 KRSLKLLSLIKWK-GSKSKLNKGFWKHLVEMPIKKK-----EMLPRCHVLDGAE----- 546
 QY 335 RGRVPEGRALDSEVPDP 352
 DB 547 -----QGLFGLDLP 557

RESULT 9
 ID AAY14128
 AAY14128 standard; Protein; 686 AA.
 AC AAY14128;
 XX 22-JUL-1999 (first entry)
 DE Human IL-1RD8 protein sequence.
 DE IL-1RD9; IL-1RD8; IL-1RD10; interleukin-1; IL-1receptor DNA; therapy;
 KM IL-1 receptor-like protein; abnormal expression; immunological disorder;
 KM inflammatory disorder; morphological disorder.
 OS Homo sapiens.
 OS WO919480-A2.
 PN WO919480-A2.
 PD 22-APR-1999.
 PF 14-OCT-1998; 98WO-US20939.
 PR 10-AUG-1998; 98US-0095987.
 PR 15-OCT-1997; 97US-0951829.
 PR 17-NOV-1997; 97US-0971635.
 PR 12-MAR-1998; 98US-0078808.
 PR 18-MAR-1998; 98US-0040714.
 PR 15-APR-1998; 98US-0081883.
 PA (SCHE) SCHERING CORP.
 PI Bazan JF, Debets JEMA, Kastelein RA, Sana TR, Timans JC;
 DR WPI; 1999-326545/27.
 DR N-PSDB; AAX58245.
 PT Interleukin-1 receptor-like polypeptides RD8, 9 and 10
 PS Claim 18; Page 93-96; 150pp; English.
 XX This sequence is a interleukin-1 (IL-1) receptor-like polypeptide,

CC designated IL-1 receptor DNAX designation 8 (IL-1RD8), of the
 CC invention. The IL-1RD8, IL-1RD9, and IL-1RD10 proteins, their fragments
 CC and mutants, also related antibodies, other binding agents and
 CC (antagonists are used to treat conditions associated with abnormal
 CC expression of the polypeptide or abnormal expression of, or response to,
 CC their ligands, e.g. immunological, inflammatory or morphological
 CC disorders. They may also be used to screen for binding agents (potential
 CC drugs), diagnostic reagents (to detect the proteins or their ligands) and
 CC to isolate related sequences (to detect the proteins or their ligands) and
 CC anti-idiotypic antibodies as carriers for toxins, radionuclides or other
 CC therapeutic agents, and for affinity purification.

XX Sequence 686 AA;

SO Query Match

Best Local Similarity 12.5%; Score 267.5; DB 20; Length 686;

Matches 103; Conservative 72; Mismatches 143; Indels 60; Gaps 19;

QY 6 DRAPFLSPEDQ--VLRPALGSSVALNCTAMV--VSGPHCSIPSYOMLK-DGLPLGIGH 61
 DB 237 DKPPRPLEPMENQPSVIDVQKPLNIPCKAFEGSGE--SGPMIYMKGEKPIEELAGH 294
 Q 62 YSLHEYSWKANLSE-----VLVSSVGVNVTSTEVYGAFTCSIONIS--FSSFTLOR 112
 DB 295 IREGIRLLKEHLGKEVELALIFDSVEADLAN-----YTCHVENRNGRKNASVLLRK 348
 QY 113 AGPTSHV--AAVLASLVLLALLALLYKCRNLVLMYODAYGEVEIN-DGKLYDAYV 169
 DB 349 KDLYKIETLAGGLGALIFLLVLLV--VIYKCYNIEMLEFYRQHFAGDETDNKRKEYDAYL 406
 QY 170 SYS-----DC--PEDRKFNVTILKPOLERRRGYKFLDPRDLPRAPESADLVNLSR 220
 DB 407 SYTKVDOTLDDCNPEEBOFALEVLDPVLEKNHYGKFLFIPERDLIPSGTYMEDLTRRYEVO 466
 QY 221 CRRLLIVYS-DAPLSRWCSHSFREGICRL--ELTRRPITFTFGORRPAHNAIRLL 276
 DB 467 SRRLLIVFTDYLIRKMSIFELSRHNMVLSGET--KVILLECTELKGKYNCOEVESL 524
 QY 277 ROHRHLYVTLMLPRGSVTP--SSDFKKEVQALPRKRVYRVEGDPQLOLDDKPMIL 334
 DB 525 KRSIKILSLIKWK-GSKSSKLNKFKHLYEPIKKK---EMLPCHVLDISAE----- 574
 QY 335 RGRVPEGRALDSEVDDP 352
 DB 575 -----QGLFGELQIP 585

RESULT 10

AA14130

ID AA14130 standard; Protein: 579 AA.

AC AA14130;

XX 22-JUL-1999 (first entry)

DE Human IL-1RD8 protein sequence.

KW IL-1RD9; IL-1RD8; IL-1RD10; Interleukin-1; IL-1receptor DNAX; therapy;
 KW IL-1 receptor-like protein; abnormal expression; immunological disorder;
 KW inflammatory disorder; morphological disorder.

OS Homo sapiens.

XX WO9919480-A2.

XX 22-APR-1999.

XX 14-OCT-1998; 98WO-US20939.

XX 10-AUG-1998; 98US-0095987.

XX 15-OCT-1997; 97US-0951829.

XX 17-NOV-1997; 97US-0971635.

XX 12-MAR-1998; 98US-0078008.

PR 18-MAR-1998; 98US-0040714.
 PR 15-APR-1998; 98US-0081883.
 XX (SCHE) SCHERING CORP.

PI Bazan JF, Debets JMA, Kastelein RA, Sana TR, Timans JC;
 DR WPI: 1999-326545/27.

DR N-PSDB; AAX58247.

XX Interleukin-1 receptor-like polypeptides RD8, 9 and 10

PS Claim 20; Page 88-91; 150pp; English.

CC This sequence is a interleukin-1 (IL-1) receptor-like polypeptide,
 CC designated IL-1 receptor DNAX designation 8 (IL-1RD8), of the
 CC invention. The IL-1RD8, IL-1RD9, and IL-1RD10 proteins, their fragments
 CC and mutants, also related antibodies, other binding agents and
 CC (ant)agonists are used to treat conditions associated with abnormal
 CC expression of the polypeptide or abnormal expression of, or response to,
 CC their ligands, e.g. immunological, inflammatory or morphological
 CC disorders. They may also be used to screen for binding agents (potential
 CC drugs), diagnostic reagents (to detect the proteins or their ligands) and
 CC to isolate related sequences (to detect the proteins or their ligands) and
 CC anti-idiotypic antibodies, as carriers for toxins, radionuclides or other
 CC therapeutic agents, and for affinity purification.

XX Sequence 579 AA;

Query Match

Best Local Similarity 12.0%; Score 257.5; DB 20; Length 579;

Matches 96; Conservative 66; Mismatches 130; Indels 45; Gaps 18;

QY 6 DRAPFLSPEDQ--VLRPALGSSVALNCTAMV--VSGPHCSIPSYOMLK-DGLPLGIGH 61
 DB 232 DKPPRPLEPMENQPSVIDVQKPLNIPCKAFEGSGE--SGPMIYMKGEKPIEELAGH 289
 QY 62 YSLHEYSWKANLSE-----VLVSSVGVNVTSTEVYGAFTCSIONIS--FSSFTLOR 112
 DB 290 IREGIRLLKEHLGKEVELALIFDSVEADLAN-----YTCHVENRNGRKNASVLLRK 343
 QY 113 AGPTSHV--AAVLASLVLLALLALLYKCRNLVLMYODAYGEVEIN-DGKLYDAYV 169
 DB 344 KDLYKIETLAGGLGALIFLLVLLV--VIYKCYNIEMLEFYRQHFAGDETDNKRKEYDAYL 401
 QY 170 SYS-----DC--PEDRKFNVTILKPOLERRRGYKFLDPRDLPRAPESADLVNLSR 219
 DB 402 SYTKVDOTLDDCNPEEBOFALEVLDPVLEKNHYGKFLFIPERDLIPSGTYMEDLTRRYE 461
 QY 220 RCRRLIVYS-DAPLSRWCSHSFREGICRL--ELTRRPITFTFGORRPAHNAIRLL 275
 DB 462 QSRRLIVFTDYLIRKMSIFELSRHNMVLSGET--KVILLECTELKGKYNCOEVES 519
 QY 276 LROHRHLYVTLMLPRGSVTP--SSDFKKEVQALPRKRVYRVEGDPQLOLDDKPMIL 310
 DB 520 LKRSIKILSLIKWK-GSKSSKLNKFKHLYEPIKK 555

RESULT 11

ABBI0347

ID ABBI0347 standard; Protein: 481 AA.

XX ABBI0347;

XX 10-JAN-2002 (first entry)

DE Human cDNA SEQ ID NO: 655.

KW Human; gene therapy; neural disorder; immune system disorder;
 KW muscular disorder; reproductive disorder; gastrointestinal disorder;
 KW pulmonary disorder; cardiovascular disorder; renal disorder;
 KW proliferative disorder; inflammation.

XX Rosen CA, Barash SC, Ruben SM;
 XX WPI: 2001-476161/51.
 DR N-PSDB: ABA06569.

PT Isolated nucleic acid molecule encoding an inflammation-associated
 PT polypeptide is used in preventing, treating or ameliorating a medical
 PT condition.

PS Claim 11: SEQ ID NO: 655; 859pp + Sequence Listing; English.

CC The present invention provides human cDNAs, proteins and related genomic
 CC DNAs. These can be used in the treatment of neural, immune system,
 CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
 CC renal and proliferative disorders and inflammation. The present sequence
 CC is a protein of the invention.

XX Sequence 481 AA:

Query Match 11.9%; Score 256.5; DB 22; Length 481;
 Local Similarity 26.5%; Pred. No. 3.9e-18;
 Matches 95; Conservative 64; Mismatches 149; Indels 51; Gaps 15;

QY 6 DRAPFLSPSEDO--VLRPALGSSVALNCTAWVYSGPHCSLPSYQWLKDGILGIGHYS 63

DB 24 DKPFLKLPMEKSLTIQETOLDGSANLTCRAFFGYSQDVS-PLIYMK-----GEKF 74

QY 64 LHEYS-----W-----VKANLSEVLVSSVLGVNVTSTEVGAFTCSIONIS---FSFTL 110

DB 75 IEDLDENRWESDIRLKEHGEDEVSTLIVDSVEEGDLGNSCYVENGNGRRHASVLL 134

QY 111 ORAG--PTSHVAVLASLVLALLLALVVKRLNVLWYDAYGEVEIN-DGKLYDA 167

DB 135 HKRELMTVELAGGAILLLVCLVT--IKCYKIEIMLFYRNHFGAEELDGKNDYDA 192

QY 168 YVSYS-----DCPEDRKVFNFILKPOLERRRGYKFLDDBDLPRAPESADLVNL 218

DB 193 YLSTYKVPDQWNOETGEERFALEILPDMLKHYKFLIPDRDLPTGTYIEDVARCV 252

QY 219 SSCRRLIVLSDAF--LSRAKCSHSFREGICRLLELTR-RPIFTFEGORRPAHPARLL 276

DB 253 DSKRLIIVMPFNIVYVRGWSIFLELTRRLNMLVTGELKVLILCSELGIMNQEVEAL 312

QY 277 ROHRHLVTLMLR-PGSVTPSSDEWKEVQALPRKVRAPV-----EGDPOTLOQ 325

DB 313 KHTIKLTVIKWHPKCKNLKSKFWKRLQYEMPER-RLEPITHQALDVSEOGPFGLQ 370

RESULT 12
 AA14129 standard; Protein: 486 AA.

XX AA14129;

XX 22-JUL-1999 (first entry)

XX Human IL-1RD10 protein sequence.

XX IL-1RD9; IL-1RD8; IL-1RD10; interleukin-1; IL-1 receptor DNAX; therapy;
 KW IL-1 receptor-like protein; abnormal expression; immunological disorder;
 XX inflammatory disorder; morphological disorder.

OS Homo sapiens.

XX W09919480-A2.

XX 22-APR-1999.

XX 14-OCT-1998; 98WO-US20939.

XX 10-AUG-1998; 98US-0095987.

XX 15-OCT-1997; 97US-0951829.

PR 17-NOV-1997; 97US-0971635.
 PR 12-MAR-1998; 98US-0078008.
 PR 18-MAR-1998; 98US-0040714.
 PR 15-APR-1998; 98US-0061883.

XX (SCHE) SCHERING CORP.

XX Bazen JF, Debets JEMA, Kastelein RA, Sana TR, Timans JC;
 DR WPI: 1999-325545/27.
 DR N-PSDB: AAX58246.

XX Interleukin-1 receptor-like polypeptides RD8, 9 and 10

PS Claim 18: Page 122-124; 150pp; English.

CC This sequence is a interleukin-1 (IL-1) receptor-like polypeptide,
 CC designated IL-1 receptor DNAX designation 10 (IL-1RD10), of the
 CC invention. The IL-1RD8, IL-1RD9, and IL-1RD10 proteins, their fragments
 CC and muteins, also related antibodies, other binding agents and
 CC (ant)agonists are used to treat conditions associated with abnormal
 CC expression of the polypeptide or abnormal expression of, or response to,
 CC their ligands, e.g. immunological, inflammatory or morphological
 CC disorders. They may also be used to screen for binding agents (potential
 CC drugs), diagnostic reagents (to detect the proteins or their ligands) and
 CC anti-idiotypic antibodies, as carriers for toxins, radionuclides or other
 CC therapeutic agents, and for affinity purification.

XX Sequence 486 AA:

Query Match 11.9%; Score 256.5; DB 20; Length 486;
 Best Local Similarity 26.5%; Pred. No. 3.9e-18;
 Matches 95; Conservative 64; Mismatches 149; Indels 51; Gaps 15;

QY 6 DRAPFLSPSEDO--VLRPALGSSVALNCTAWVYSGPHCSLPSYQWLKDGILGIGHYS 63

DB 29 DKPFLKLPMEKSLTIQETOLDGSANLTCRAFFGYSQDVS-PLIYMK-----GEKF 79

QY 64 LHEYS-----W-----VKANLSEVLVSSVLGVNVTSTEVGAFTCSIONIS---FSFTL 110

DB 80 IEDLDENRWESDIRLKEHGEDEVSTLIVDSVEEGDLGNSCYVENGNGRRHASVLL 139

QY 111 ORAG--PTSHVAVLASLVLALLLALVVKRLNVLWYDAYGEVEIN-DGKLYDA 167

DB 140 HKRELMTVELAGGAILLLVCLVT--IKCYKIEIMLFYRNHFGAEELDGKNDYDA 197

QY 168 YVSYS-----DCPEDRKVFNFILKPOLERRRGYKFLDDBDLPRAPESADLVNL 218

DB 198 YLSTYKVPDQWNOETGEERFALEILPDMLKHYKFLIPDRDLPTGTYIEDVARCV 257

QY 219 SSCRRLIVLSDAF--LSRAKCSHSFREGICRLLELTR-RPIFTFEGORRPAHPARLL 276

DB 258 DSKRLIIVMPFNIVYVRGWSIFLELTRRLNMLVTGELKVLILCSELGIMNQEVEAL 317

QY 277 ROHRHLVTLMLR-PGSVTPSSDEWKEVQALPRKVRAPV-----EGDPOTLOQ 325

DB 318 KHTIKLTVIKWHPKCKNLKSKFWKRLQYEMPER-RLEPITHQALDVSEOGPFGLQ 375

RESULT 13
 AA195298 standard; Protein: 696 AA.

XX AA195298;

XX 12-SEP-2000 (first entry)

XX Human Xrec2.

XX Xrec2; interleukin-1 receptor; human; therapy; inflammation; fever.

XX Homo sapiens.

XX MO200036108-A2.
PN
XX 22-JUN-2000.
PD
XX 14-DEC-1999; 99WO-US29549.
XX
XX 14-DEC-1998; 98US-0112163.
XX 10-NOV-1999; 99US-0164675.
XX (IMMUNEX CORP.
PA
XX Sims JE, Smith DE, Born TL;
PI
XX MPI: 2000-442387/38.
DR N-PSDB; AAA27919.
DR
XX
PT Isolated interleukin-1 (IL-1) zeta nucleic acids and splice variants
PT ID21, ID22, ID23 and their encoding proteins, useful as probes for
PT identifying genes associated with diseases such as glaucoma, and
PT insulin-dependent diabetes mellitus
XX
XX Claim 21; Page 11; 87pp; English.
XX
XX The present sequence is that of human Xrec2, a member of the
XX interleukin-1 (IL-1) receptor family. The sequence was determined
XX by translation of the nucleotide sequence of isolated IL-1 zeta
XX cDNA (see AAA27919). The invention is directed to novel, purified
XX and isolated IL-1 zeta, IL-1 zeta splice variants (see AAA27918-22)
XX polypeptides (see AAY95297-301), the nucleic acids (see AAA27918-22)
XX encoding such polypeptides, processes for production of recombinant
XX forms of such polypeptides, and their uses. The polypeptides can
XX be used to study cellular processes such as immune regulation, cell
XX proliferation, cell death, cell migration, cell-to-cell interaction
XX and inflammatory responses, to identify proteins associated with
XX IL-1 zeta, to screen for potential inhibitors, and to prepare
XX antibodies. In particular, they can be used to activate and/or
XX inhibit the activation of vascular endothelial cells and
XX lymphocytes, induce and/or inhibit the induction of local tissue
XX destruction and fever, inhibit and/or stimulate macrophages and
XX vascular endothelial cells to produce IL-6, induce and/or inhibit
XX the induction of prostaglandins, nitric oxide synthetase, and
XX metalloproteinases, and upregulate and/or inhibit the upregulation of
XX molecules on the surface of vascular endothelial cells.
XX
XX Sequence 696 AA;
SO
XX
XX Query Match 11.9%; Score 256.5; DB 21; Length 696;
XX Best Local Similarity 26.5%; Pred. No. 6.9e-18;
XX Matches 95; Conservative 64; Mismatches 149; Indels 51; Gaps 15;

DB 528 KHTIKLTVIRKMGPKCNKLSNFKWRLQYEMPFK-RIEPIITHEQALDVSQGPGEIQ 585
RESULT 14
ID AAM78808 standard; Protein; 696 AA.
XX
XX AAM78808;
AC
XX 06-NOV-2001 (first entry)
DT
XX
XX Human protein SEQ ID NO 1470.
DE
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
XX Homo sapiens.
OS
XX
XX WO200157190-A2.
EN
XX
XX 09-AUG-2001.
PD
XX
XX 05-FEB-2001; 2001WO-US04098.
PF
XX
XX 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0683325.
PR 30-NOV-2000; 2000US-0728422.
XX
XX (HYSEQ-) HYSEQ INC.
PA
XX
XX Tang YT, Liu C, Dymnac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue A, Yang Y, Wejhrman T, Goodrich R;
XX
XX MPI: 2001-476283/51.
DR N-PSDB; AAK51941.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
PT
XX
XX Claim 20; Page 3744-3746; 6221pp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
XX encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation.
XX Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
XX (AAM80020) are omitted as the relevant pages from the sequence listing
XX were missing at the time of publication.
XX
XX Sequence 696 AA;
SQ
XX
XX Query Match 11.9%; Score 256.5; DB 22; Length 696;
XX Best Local Similarity 26.5%; Pred. No. 6.9e-18;
XX Matches 95; Conservative 64; Mismatches 149; Indels 51; Gaps 15;

```

OY 64 LHEYS-----W-----VKANLSEVLVSSVGLVNTSTREVGATCSTIONIS---FSSFTL 110
DB 290 IEDDENRWESDRIKILKEHIGDEOVSTISLVDSEEGDIGNTSCYVNGNGRRHRAVSL 349
OY 111 QRAG--PTSHVAAYLASLVLALLALLAALLYVKCRNLVLLMYODAYGEVEIN-DGKLYDA 167
DB 350 HKRELMTYVELAGGAGALLLVCLVLT--IYCKYKIEIMLFYRNHFGAEELDDNDYDA 407
OY 168 YVSYS-----DCEPDREKVFNFILKPOLERRRGYKFLDPRDLPRAPESADLLVNL 218
DB 408 YLSYTKVPDQNMDETGEERFALEILPDMLEKHYGKFLPDRDLPTGTYTEDEVARCV 467
OY 219 SSCRRLIYVLSDAF--LSRAMCSHSFREGICRLLELTR-RPIFTFEQORDPAHPALRL 276
DB 468 DQSKRLIIVMTPNVYVRRGWSIFELETRLRNMLVTGEEKVILLIECSBELRGIMNYQVEAL 527
OY 277 ROHRRLVTLMLR--PGSVTPSSDPEWKEVOLAPRKVRYRPY-----EGDPQTOLO 325
DB 528 KHTIKLIVIKWHGPKCNKLNKSKFMKRLQYEMPRK-RIEPTHRQALDVSEOGPFGELO 585

RI 15
AB 185
ID ABB11785 standard; peptide: 710 AA.
AC ABB11785;
XX 11-JAN-2002 (first entry)
DE Human oligophrenin-4 homologue, SEQ ID NO:2155.
XX Human: cytokine; cell proliferation; cell differentiation; growth factor;
XX haematopoiesis regulation; tissue growth; immunomodulator; activin;
XX inhibin; chemotaxis; chemokinesis; cancer; tumour; haematopoietic disorder;
XX myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
XX chronic inflammatory condition; proliferative retinopathy;
XX atherosclerosis; coronary heart disease; arterial ischaemia;
XX bone disorder; osteoporosis; vascular growth disorder;
XX tissue regeneration; wound healing; infection; immune disorder;
XX cell culture; drug screening; gene therapy; antiinflammatory;
XX antihemetic; antiarthritis; haemostatic; antihypertensive;
XX cytoskeletal; osteopathic; vasotropic; cardiac; virulence; antibacterial;
XX antifungal; vulnereary; antituber.
XX OS Homo sapiens.
XX PM WO200157188-A2.
XX 9-AUG-2001.
PE 05-FEB-2001; 2001WO-US03800.
XX 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Drmanac RT;
XX MPI: 2001-457740/49.
DR N-PSDB: ABA09029.
XX Human proteins and DNA encoding sequences useful for preventing,
XX creating or ameliorating a medical condition in a mammalian subject
XX e.g. arthritis and cancer -
XX Claim 20; Page 248; 1963pp; English.
XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
XX sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
XX invention also relates to vectors and recombinant host cells comprising a

```

```

CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation regulatory activity; stem cell growth factor activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC The invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a novel human
CC polypeptide of the invention.
XX
SQ Sequence 710 AA;
Query Match 11.9%; Score 256.5; DB 22; Length 710;
Best Local Similarity 26.5%; Pred. No. 7,1e-18;
Matches 95; Conservative 64; Mismatches 149; Indels 51; Gaps 15;
OY 6 DRAPFLSPSPDO--VLRALGSSVALNCRAMVYSGHCLSPVOMLKDPLGICGHHS 63
DB 253 DKPRKLLYPMESKLIQTQIGDSANLTCRAFRGYSADV-PLIYMK-----GEKF 303
OY 64 LHEYS-----W-----VKANLSEVLVSSVGLVNTSTREVGATCSTIONIS---FSSFTL 110
DB 304 IEDDENRWESDRIKILKEHIGDEOVSTISLVDSEEGDIGNTSCYVNGNGRRHRAVSL 363
OY 111 QRAG--PTSHVAAYLASLVLALLAALLYVKCRNLVLLMYODAYGEVEIN-DGKLYDA 167
DB 364 HKRELMTYVELAGGAGALLLVCLVLT--IYCKYKIEIMLFYRNHFGAEELDDNDYDA 421
OY 168 YVSYS-----DCEPDREKVFNFILKPOLERRRGYKFLDPRDLPRAPESADLLVNL 218
DB 422 YLSYTKVPDQNMDETGEERFALEILPDMLEKHYGKFLPDRDLPTGTYTEDEVARCV 481
OY 219 SSCRRLIYVLSDAF--LSRAMCSHSFREGICRLLELTR-RPIFTFEQORDPAHPALRL 276
DB 482 DQSKRLIIVMTPNVYVRRGWSIFELETRLRNMLVTGEEKVILLIECSBELRGIMNYQVEAL 541
OY 277 ROHRRLVTLMLR--PGSVTPSSDPEWKEVOLAPRKVRYRPY-----EGDPQTOLO 325
DB 542 KHTIKLIVIKWHGPKCNKLNKSKFMKRLQYEMPRK-RIEPTHRQALDVSEOGPFGELO 599

```

Search completed: November 7, 2002, 18:49:32
Job time : 66 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2002, 18:42:16 ; Search time 70 Seconds
(without alignments)
1013.256 Million cell updates/sec

Title: US-09-598-443-2

Perfect score: 2147

Sequence: 1 MPGVCDRAPEFLSPSEDOVL.....GSRNYSARTDFYCLVSKDDM 410

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Sequenced: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MHC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_PROTOZOA:*
12: SP_VIRUS:*
13: SP_VERTebrATE:*
14: SP_UNCLASSIFIED:*
15: SP_VIRUS:*
16: SP_BACTERIA:*
17: SP_ARCHAEA:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2147	100.0	410	Q9H733	Q9H733 homo sapien
2	1546.5	72.0	409	Q9JLZ8	Q9JLZ8 mus musculu
3	267.5	12.5	658	Q9NZNO	Q9NZNO homo sapien
4	267.5	12.5	686	Q9NP60	Q9NP60 homo sapien
5	261	12.2	686	Q9ERS6	Q9ERS6 mus musculu
6	256.5	11.9	696	Q9UJ53	Q9UJ53 homo sapien
7	256.5	11.9	696	Q9UJ53	Q9UJ53 homo sapien
8	254	11.8	537	Q61098	Q61098 mus musculu
9	231.5	10.8	541	Q13478	Q13478 homo sapien
10	228.5	10.6	570	Q14915	Q14915 homo sapien
11	217	10.1	574	Q14915	Q14915 mus musculu
12	215.5	10.0	570	Q61730	Q61730 mus musculu
13	211	9.8	533	Q9DEE1	Q9DEE1 galus galli
14	210.5	9.8	570	Q63621	Q63621 ratuS norv
15	210	9.8	533	Q9DEE5	Q9DEE5 galus galli
16	206.5	9.6	575	Q9HB29	Q9HB29 homo sapien

17	204.5	9.5	561	11	Q62929	Q62929 ratuS norv
18	204	9.5	562	4	Q13525	Q13525 homo sapien
19	203.5	9.5	573	6	Q9TV71	Q9TV71 equuS cabal
20	202	9.4	556	4	Q9UQ44	Q9UQ44 homo sapien
21	196.5	9.2	1346	5	Q9V477	Q9V477 drosophila
22	189.5	8.8	567	11	Q05208	Q05208 mus musculu
23	189	8.8	599	4	Q95256	Q95256 homo sapien
24	184	8.6	555	13	Q90874	Q90874 galus galli
25	183	8.5	566	11	Q62611	Q62611 ratuS norv
26	177.5	8.3	614	11	Q92281	Q92281 mus musculu
27	161	7.5	1100	5	Q24622	Q24622 drosophila
28	152.5	7.1	795	5	Q9VJX9	Q9VJX9 drosophila
29	152.5	7.1	795	5	Q9NBK8	Q9NBK8 drosophila
30	150.5	7.0	743	5	Q95RV9	Q95RV9 drosophila
31	150.5	7.0	795	5	Q9XZF9	Q9XZF9 drosophila
32	150.5	7.0	1514	5	Q9VUN0	Q9VUN0 drosophila
33	150.5	7.0	1514	5	Q9NBK9	Q9NBK9 drosophila
34	149.5	7.0	557	5	Q961H0	Q961H0 drosophila
35	149.5	7.0	1385	5	Q9V8Z5	Q9V8Z5 drosophila
36	145.5	6.8	486	13	Q9DEE2	Q9DEE2 galus galli
37	144.5	6.7	486	13	Q9DEE4	Q9DEE4 galus galli
38	144	6.7	784	11	Q9QDN7	Q9QDN7 mus musculu
39	144	6.7	784	11	Q9DBC4	Q9DBC4 mus musculu
40	142.5	6.6	826	6	Q9TSP2	Q9TSP2 papio anub1
41	139.5	6.5	1389	5	Q24591	Q24591 drosophila
42	138	6.4	752	4	Q9UK78	Q9UK78 homo sapien
43	138	6.4	799	4	Q9UM57	Q9UM57 homo sapien
44	138	6.4	839	4	Q00206	Q00206 homo sapien
45	137.5	6.4	1446	5	Q9V820	Q9V820 drosophila

ALIGNMENTS

RESULT 1	
Q9H733	PRELIMINARY: PRT: 410 AA.
AC Q9H733:	
DT 01-MAR-2001 (TREMBLrel. 16, Created)	
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)	
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)	
DE CDNA: FJ21446 FIS, CLONE COL04458.	
OS Homo sapiens (Human).	
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	
OX NCBI_TaxID=9606;	
ON [1]	
RP SEQUENCE FROM N.A.	
RC TISSUE=COLON;	
RA Kawabata A., Hiki T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,	
RA Okitani R., Ota T., Suzuki Y., Odayashi M., Nishi T., Shibahara T.,	
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.:	
RT "NEDO human cDNA sequencing project."	
RL Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.	
DR EMBL: AK025099; BAB15066.1; -	
DR InterPro: IPR003600; Ig_Like.	
DR InterPro: IPR000157; TIR.	
DR PIRam: PF01582; TIR: 1.	
DR SMART: SM00410; Ig_Like; 1.	
DR SMART: SM00255; TIR: 1.	
SQ SEQUENCE 410 AA; 45707 MW; 2A7A663D79567ED6 CRC64:	
Query Match	100.0%; Score 2147; DB 4; Length 410;
Best Local Similarity	100.0%; Pred. No. 2.8e-197;
Matches 410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 1 MPVCDRAPEFLSPSEDOVLPAALGSSVALNCAWVSGPHCSLPVOMLKDLPLGIGG 60	
DB 1 MPVCDRAPEFLSPSEDOVLPAALGSSVALNCAWVSGPHCSLPVOMLKDLPLGIGG 60	
QY 61 HVSLEHYSWVKANLSEVLVSVLVGVNTSTEVYGAFTCSIONISFSSFTLQRAQPTSHA 120	
DB 61 HVSLEHYSWVKANLSEVLVSVLVGVNTSTEVYGAFTCSIONISFSSFTLQRAQPTSHA 120	

QY	121	AVIASLITVLTALLLAALYYVCRLNVLWMMQDAPYGEVEINDDCKIXDAYVYSDDCEGRKF	180
Db	121	AVIASLITVLTALLLAALYYVCRLNVLWMMQDAPYGEVEINDDCKIXDAYVYSDDCEGRKF	180
QY	181	VNFILTKPOLERRRGYKLEFDRODRLPAPEBSADLVNISRCRRLTYVLSDAFISRAMCSH	240
Db	181	VNFILTKPOLERRRGYKLEFDRODRLPAPEBSADLVNISRCRRLTYVLSDAFISRAMCSH	240
QY	241	SFPEGICRLLELTRRPITFTFEGSORRPPAPALRLRQHRNHYVTLMLRPGSYPPSSDFW	300
Db	241	SFPEGICRLLELTRRPITFTFEGSORRPPAPALRLRQHRNHYVTLMLRPGSYPPSSDFW	300
QY	301	KEYOVALPKRYVRYRVEDDPTOLQDDKDMILIRGVRVSGALDSVEVDPPEBGDLGYRG	360
Db	301	KEYOVALPKRYVRYRVEDDPTOLQDDKDMILIRGVRVSGALDSVEVDPPEBGDLGYRG	360
QY	361	PVYGEBSAAPHTHSYVLSFESNSSEVDVSDGSRNYSARITDFCYLKSDDM	410
Db	361	PVYGEBSAAPHTHSYVLSFESNSSEVDVSDGSRNYSARITDFCYLKSDDM	410

RA	FA	2	09JL28	PRELIMINARY:	PRT:	409	AA.
AC	09JL28						
PT	01-OCT-2000	(TREMblrel.	15	Created)			
PT	01-OCT-2000	(TREMblrel.	15	Last sequence update)			
PT	01-DEC-2001	(TREMblrel.	19	Last annotation update)			
DE	TOLL/INTERLEUKIN-1	RECEPTOR	8.				
CN	SIGIRR OR TIR8.						
OS	Mus musculus (Mouse).						
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus						
NC	NCBI_TaxID=10090;						

RP SEQUENCE FROM N.A.
RA- Penton-Rol G., Polentarcutli N., Muzio M., Mantovani A.;
RT "TIR8: a novel toll/interleukin-1 receptor family member."
RL Submitted (DEC-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF113795; AAF62200.1; -
DR MGD: MGI:1344402; Slgflr.
DR InterPro: IPR003599; IG.
DR InterPro: IPR003006; IG_MHC.
DR InterPro: IPR000157; TIR.
DR Pfam: PF00047; Ig_1.
DR Pfam: PF01582; TIR; 1.
DR SMART: SM00409; IG; 1.
DR SMART: SM00255; TIR; 1.
DR Receptor. 409 AA; 45707 MW; DC4AB99A72356A13 CRC64;
DR SEQUENCE

Query Match	72.0%;	Score 1546.5;	DB 11;	Length 409;
Best Local Similarity	72.7%;	Pred. NO. 1.1e-139;		
Matches 303; Conservative	32;	Mismatches 67;	Indels 15;	Gaps 4

QY	1	MPVCORADDFLSPSEDOVLPRALGSSVLTANCAMVVSQPHCSLETSYOMLKDQPLKIG	60
QY	1	MAVCYMAAPNPLSPSEDOALTALGREALVNTCAVAFSPQCPQSYOMLKDGLATLNGS	60
QY	61	HSYLHSEYVWKNLSEVYVSVLGVNTSTREYVGAFTCSIONISPSFTLTOIRAQPTSHA	120
QY	61	HFLSHEDFVWVANSFSEI--VSSVLVNLNTMNEIDGTFTCSVMVSSHSTTLKRAAPACHVA	119
QY	121	AVTASLTVLTLTALLTALLTYKCRGLVNTLWQDYGXGEVLTNDCKLYDAYVSSDCEPERK	180
QY	120	AVTASLTVLTVLTVLTVLTVLTVCRNLMLMLQDITYREVENMDCKLDAVYVSSDCEPERK	179
QY	181	VNVLTKPQLERRRKYKFLFLDRDLTPRAEFSADLVNYSRCRLLTVLSDAFSLRAMCSH	240
QY	180	VNVLTKPQLERRRKYKFLFLDRDLTPRAADSADLVNYSRCRLLTVLSDAFSLRPMCSO	239
QY	241	SPDESGICRL--LDELTRRLPTTEGORDPAHPALRL--LRQHRHLVTLTLLMRGSV	293

[illegible]

RESULT 3		
Q9NZNO		
ID	Q9NZNO	PRELIMINARY: PRT; 658 AA.
AC	Q9NZNO	
DT	01-OCT-2000 (Tremblrel. 15, Created)	
DT	01-OCT-2000 (Tremblrel. 15, Last sequence update)	
DT	01-DEC-2001 (Tremblrel. 19, Last annotation update)	
DE	X-LINKED INTERLEUKIN-1 RECEPTOR ACCESSORY PROTEIN-LIKE 2	

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[]
RN SOURCE: FROM N.A.
RP MEDLINE=20218555; PubMed=10757639.
RA Jin H., Vismesvaralan R., Gardner R.J., Roberts R.G.;
RT 21 novel members of the interleukin-1 receptor gene family, one
RT deleted in Xp21.3-Xp21.3 mental retardation.";
RT Eur. J. Hum. Genet. 8:87-94(2000).

DR	InterPro: IP0003006; IG_MHC.
DR	InterPro: IP0000157; TIR.
DR	Pfam: PF00047; 15; 3.
DR	Pfam: PF01582; TIR; 1.
DR	SMART: SM00409; IG; 2.
DR	SMART: SM00410; IG-like; 1.
DR	SMART: SM00255; TIR; 1.
KW	Receptor.
PT	
QO	NON_TER 1
SEQUENCE	658 AA; 75626 MW; 5B7506AF1110BB12 CRC64;

Query Match	12.58;	Score 267.5;	DB 4;	Length 658;
Best Local Similarity	27.28;	Pred. No. 9.3e-17;		
Matches 103; Conservative	72;	Mismatches 143;	Indels 60;	Gaps 19

QY 6 DRAPDPLSSSEQO--VLPPALGSSVALNCAAW--VSGPHCSLPVSQMLK--DGLPLGIGGH 61

Db 209 DKPKPLFPFEMENQSPVIDVOLGKPLNIPCAAFGSGSG--SGPMIYMKGKFLTEELAGH 266

QY 62 YSDHEYSWKANLISE-----VLSSVLGVNVSSTVEYGAFCISQIONS-----FSSFTLOR 112

Db 267 IRGEGRRLRLEKHGEKEVEELALIPDSVVEADLAN-----YCHQENNRGKRKHAASVLRK 320

QY 113 AGPTSHV--AAVLAVLIVLALLLALALYKCLINLWYQAYEVEIN--DGLYLDAYV 169

Db 321 KDLIYKIELAGGIGAIFFLLVILV--VIYCYNIEMILFYP RHQFADGETNDNKNEYDAYL 378

QY 170 SYS-----DC--PEDRKFNPNFLIKPOLERRRGYALFLFDDDDLLPRAPEASADLLVLSR 220

Db 379 SYKRVQDDPLDQDNEEOGFLEVLPRVLEKHYGLYLFPEEDLLIPSGTYMEDITRIVEQ 438

QY 221 CRLLIIVLS--DAFLSRAMCSHSFREGDCRL--ELTRRPIPIYTEGGRRPAPHAIRLRL 276

Db 439 SRRLIIVLPDIYLRGMSIFELFSRILHNMLVSGEL--KVILIECTELKGKGVNCOEVEST 496

QY 277 RQRHRLVTLTLIRPSSVTP--SSDFWKEVQLALPRVRYRPRVEGSPQIOLDDDDKDPAILL 334

```

Db 497 KRSTLSTLJMK-6SKSKLNSKRMKHLVEMPIKK-----EMLRCHVDSAE----- 546
Oy 335 RGRNPEGRALDSEVDDP 352
Db 547 -----OGLFGELOPIP 557

RESULT 4
O9NP60 PRELIMINARY; PRT; 686 AA.
O9NP60:
AC 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 15, last annotation update)
DE ILIRAPL-2 RELATED PROTEIN (INTERLEUKIN-1 RECEPTOR 9) (IL-1 RECEPTOR ACCESSORY PROTEIN-LIKE 2) (TIGIR-1).
GN ILIRAPL-2 OR ILIR9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI.TaxID=9606;
RN 1
RP SEQUENCE FROM N.A.
RC TISSUE=ADULT BRAIN;
RA Ferrante M.;
RL Submitted (May-2000) to the EMBL/GenBank/DBJ databases.
RN 2
RP SEQUENCE FROM N.A.
RA Sana T.R., Debets R., Timans J.C., Bazan J.F., Kastalein R.A.;
RT "Computational identification, cloning and characterization of IL-1R9, a novel interleukin-1 receptor (IL-1R)-like gene encoded on human chromosome Xq 22.2-22.3."
RT - Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN 3
RP SEQUENCE FROM N.A.
RC TISSUE=BRIN;
RA Grabowski M., Lorenz B., Hubel R., Strom T.M.;
RT "A gene (ILIRAPL-2) with 61% identity to ILIRAPL maps to Xq22.2."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN 4
RP SEQUENCE FROM N.A.
RX MEDLINE=20459050; PubMed=10882729;
RA Born T.L., Smith D.E., Garza K.E., Renshaw B.R., Bertles J.S., Slins J.E.;
RT "Identification and characterization of two members of a novel class of the interleukin-1 receptor (IL-1R) family. Delineation of a new class of IL-1R-related proteins based on signaling."
RN 5
RA EMBL; AJ290436; CAB89867.1; -
DR EMBL; AF212016; AAF61307.1; -
DR EMBL; AJ272208; CAB86868.1; -
DR EMBL; AF284436; AAC21370.1; -
DR InterPro: IPR003599; IG_
DR InterPro: IPR003600; IG_1ike.
DR InterPro: IPR003006; IG_MHC.
DR InterPro: IPR004075; Intriknl_receptor1.
DR InterPro: IPR000157; TIR.
DR Pfam: PF00047; IG_3.
DR Pfam: PF01582; TIR; 1.
DR PRINTS; PRO1537; INTRLNRI1F.
DR SMART; SM00409; IG; 2.
DR SMART; SM00410; IG_1ike; 1.
DR SMART; SM00255; TIR; 1.
KW Receptor.
SO SEQUENCE 686 AA; 78669 MW; E400F7ECD186957C CAC64;

Query Match 12.5%; Score 267.5; DB 4; Length 686;
Best Local Similarity 27.2%; Pred. No. 9.9e-17;
Matches 103; Conservative 72; Mismatches 143; Indels 60; Gaps 19;
Oy 6 DRAPDFLSPSEDQ--VLPRALGSSVALNCTAMV--VSGPRCSIPSPQWLK-DCLPLGICGH 61
Db 237 DKPPPLPFPNPNPSVIDVQLKPLNICKAFPFSGE--SGPMIYWMKGKFFIEELAGH 294

```

```

QY 62 YSLHEYSMWKNIKE-----VLVSVLGVAVSTEEYGAFTCSIONIS---FSSFLOR 112
Db 295 IRBEIRIKLRKHLEKEVELALITDSVYENDLAN-----YTCHEVRNRKRKIASVYLRK 348
QY 113 AGPPISHV--AAVLASLVLALLLALLLALLYKCRNLVMTLQDAYGEVEIN-DGKLYDAYV 169
Db 349 KDLIYKIELAGLGAIFLLLVLLV--VIYCYUNIELMLFYRHHGAGDETMDNKREYDAYL 406
QY 170 SYS-----DC--PEDRKFNVLFLKPOLERRRGYKFLDDRDLLPRAEPSADLVNLSR 220
Db 407 SYTRVDDQDITLDCDNRPEEBOFLAEVLPVYLKHKYKLFIEPDRDILPSGYTMMEDLTRVEQ 466
QY 221 CRLLIVVLIS-DAFISRAMCISFEFGICRLL--ELTRPPIFTTFESQRRDPAPHALRL 276
Db 467 SRRLIIVLTPDYILIRKMSIFELERSLHNNLVSGEI--KVILIECTELKRVKCOQEVESL 524
QY 277 ROHHLVLLTLLMRGQSYTP--SSDFKREVOVLALPRKRVYRPVGGDPQYOLQDKDRPMLIL 334
Db 525 KRSLIKLLSLIKWK-GSKMSKSKNSKFMGHIVLYEMVITKKK---EMLIPRCHILDSAE----- 574
QY 335 RGRVPEGRALDSEVDPDP 352
Db 575 -----QGELGELQIP 585

```

RESULT 5			
Q9ERS6	Q9ERS6	PRELIMINARY;	PRF; 686 AA.
AC	Q9ERS6;		
DT	01-MAR-2001 (Tremblrel, 16, Created)		
DT	01-MAR-2001 (Tremblrel, 16, Last sequence update)		
DT	01-DEC-2001 (Tremblrel, 19, Last annotation update)		
DE	TLGIR-1.		
CN	ILIRAPL2.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=204559050; PubMed=10882729;		
RA	Born T.L., Smith D.E., Garika K.E., Renshaw B.R., Bertles J.S.,		
RA	Sims J.E.;		
RT	"Identification and characterization of two members of a novel class		
RT	of the interleukin-1 receptor (IL-1R) family. Delineation of a new		
RT	class of IL-1R-related proteins based on signaling.";		
RL	J. Biol. Chem. 275:29946-29954(2000).		
DR	EMBL; AF284437; AAC31371.1; -		
DR	MGD; MGI:1913106; Il1rapl2.		
DR	InterPro; IPR003599; IG.		
DR	InterPro; IPR003600; IG_like.		
DR	InterPro; IPR003006; IG_MHC.		
DR	InterPro; IPR000157; TIR.		
DR	Pfam; PF00047; 19; 3.		
DR	Pfam; PF01582; TIR; 1.		
DR	SMART; SM00409; IG; 2.		
DR	SMART; SM00410; IG_like; 2.		
DR	SMART; SM00255; TIR; 1.		
DR	SEQUENCE 686 AA; 78797 MW; 36160D1CDE9B8264 CRC64;		
QY	Query Match	12.2%; Score 261; DB 11; Length 686;	
Db	Best Local Similarity	28.9%; Pred. No. 4.2e-16;	
Matches	98; Conservative	65; Mismatches 126; Indels	50; Gaps
QY	6 DRAPDELSPSEDQ--VLRPALGSSVALNCTAMV-VSGPHCSLPYQWLK-DQLPLIGIGH 61		
Db	237 DKPXPPLPEPMENQPSVIDVQGLKPLINPCAKAFGSGE--SGPMIYMKKGKFIIBELAGH 294		
OY	62 KSLHEYSWVKANLSE-----VLVSSYGLGVANTSTIEVYGAFICSLQINIS---FSSFTLQIR 112		
Db	295 IREGRIILRLKKEKEVELTLIFGFSYAADLAN-----YTCHEVNNNGRKHAGVLLRK 348		

```

OY 113 AGPSTHV--AAVLASLVLALLAALLVYKRLNVLWYODAYGEVE-INDKRLYDAYV 169
DB 349 KDLTYIELAGLIGALIFLLILLVLT--VYKCYNIEMLEFYORFGDETDONKNEYDAYL 406
OY 170 SYS-----DC-PEDRKFNFIILKPOLERRRGYKFLDRLDRLPRAEPSADLLVNLNR 220
DB 407 SYTKVDQDTLDCDNTNEEDQFALEILPDVLEKHGYKFLIPERDIL---PSGTYIEDLTR 462
OY 221 C---RLIVLS-DALFSRAWCSHFREGLCRL---ELTRRPFTTFEGQRDPANPA 272
DB 463 CYESGRRLIVLTPYILIRGWSIFELSRNLHNMVSGEI--KVILLECTELKGVNCOE 520
OY 273 LRLRQHRILVTLMLR-PSGVTSSDFWKEVQALPCK 310
DB 521 VESLKHNIKLLSLIKWKGPKSSKLSKFWKHLYEMPIK 559

RESULT 6
ID Q90J53 PRELIMINARY; PRT: 696 AA.
AC Q90J53.
DB 01-MAY-2000 (Tremblrel. 13, Created)
DB 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DB 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE OLIGOEPHENIN-4 (TIGIR-2).
GN OPN4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99400548; PubMed=10471494;
RA Carle A., Jun L., Bienvenu T., Vinet M.;
RT "A new member of the IL-1 receptor family highly expressed in
RT hippocampus and involved in X-linked mental retardation.";
RN [2]. Genet. 23:25-31(1999).
RP SEQUENCE FROM N.A.
RX MEDLINE=20459050; PubMed=10882729;
RA Born T., Smith D.E., Garika K.E., Renshaw B.R., Bertles J.S.,
RT "Identification and characterization of two members of a novel class
RT of the interleukin-1 receptor (IL-1R) family. Delineation of a new
RT class of IL-1R-related proteins based on signaling.";
RL J. Biol. Chem. 275:29946-29954(2000).
DR EMBL: AJ243874; CABS6046-29954(2000).
DR EMBL: AF284435; CAB21369.1; -.
DR InterPro: IPR003598; 19-C2.
DR InterPro: IPR003006; 19-C2.
DR InterPro: IPR004075; 19-C2.
DR InterPro: IPR000770; SAND.
DR InterPro: IPR000157; TIR.
DR Pfam: PF00047; 197_2.
DR Pfam: PF01582; TIR.
DR PRINTS: PRO1537; INTRILKINR1F.
DR SMART: SM00408; ITC2; 1.
DR SMART: SM00258; SAND; 1.
DR SMART: SM00255; TIR; 1.
KM Immunoglobulin domain.
SQ SEQUENCE 696 AA; 79968 MW; 987ADB503D73C9A9 CRC64;

Query Match 11.9%; Score 256.5; DB 4; Length 696;
Best Local Similarity 26.5%; Pred. NO. 1.1e-15;
Matches 95; Conservative 64; Mismatches 149; Indels 51; Gaps 15;

```

```

OY 111 ORAG--PTSHVAVALSLVLLALLAALLVYKRLNVLWYODAYGEVEIN-DGRLYDA 167
DB 350 HKRELMTYIELAGGAILLVLCLVLT--IKCYKIEIMLEYRHHNFAEELDGKNDKYDA 407
OY 168 YSYS-----DCPEDRKFNFIILKPOLERRRGYKFLDRLDRLPRAEPSADLLVNL 218
DB 408 YLSYTKVDPDQNNQETGEERFALEILPDMLEKHGYKFLIPDRDLIPGTGYIEDVAVCY 467
OY 219 SECRRLIVYLSDAF-LSRAWCSHFREGLCRLLETR-RPIFTTFEGQRDPANPALRL 276
DB 468 DSKRLIVLTPYILIRGWSIFELSRNLHNMVSGEI--KVILLECTELKGVNCOE 520
OY 277 RQHRILVTLMLR-PSGVTSSDFWKEVQALPCKVRYRYV-----EGDPOTOLQ 325
DB 528 KHTKLLTVIRMHGPKCNKLSKFWKRLQYEMPER-RIEPTIRQALDYSEOGPFGELO 585

RESULT 7
ID Q9NZN1 PRELIMINARY; PRT: 696 AA.
AC Q9NZN1.
DB 01-OCT-2000 (Tremblrel. 15, Created)
DB 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DB 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE X-LINKED INTERLEUKIN-1 RECEPTOR ACCESSORY PROTEIN-Like 1.
GN ILIRKRL1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20218565; PubMed=10757639;
RA Jin H., Wisnesvaran R., Gardner R.J., Roberts R.G.;
RT "Two novel members of the interleukin-1 receptor gene family, one
RT deleted in human XP2.3 mental retardation.";
RL Eur. J. Hum. Genet. 8:87-94(2000).
DR EMBL: AF181284; CABS5941.1; -.
DR InterPro: IPR003598; 19-C2.
DR InterPro: IPR003006; 19-C2.
DR InterPro: IPR004075; 19-C2.
DR InterPro: IPR000770; SAND.
DR InterPro: IPR000157; TIR.
DR Pfam: PF00047; 197_2.
DR Pfam: PF01582; TIR.
DR PRINTS: PRO1537; INTRILKINR1F.
DR SMART: SM00408; ITC2; 1.
DR SMART: SM00258; SAND; 1.
DR SMART: SM00255; TIR; 1.
KM Immunoglobulin domain. Receptor.
SQ SEQUENCE 696 AA; 79979 MW; F2BEC371537F9AA0 CRC64;

Query Match 11.9%; Score 256.5; DB 4; Length 696;
Best Local Similarity 26.5%; Pred. NO. 1.1e-15;
Matches 95; Conservative 64; Mismatches 149; Indels 51; Gaps 15;

```

OY 219 SRCRLIYVLSDAF-LSRAMCSHSFREGJCRLELTR-RPIETFGQRDPAPALRL 276
 DB 468 DSKRLIYVTPYVVRGMSIFELETRLRNMLVTGEIVILLIECSLGINMYQVEAL 527
 OY 277 RQHRHLVTLIMR-PGSVPSSDFMKEVQALPRKVRYPV-----EGPQPOLQ 325
 DB 528 KATIKLITVIMHGPCKNLKNSFKWRLQYEMFER-RIBETTHEQALDVSEQGFGLQ 585

RESULT 8

O61098 PRELIMINARY; PRT: 537 AA.
 AC 061098;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE IL-1RIP PRECURSOR.
 GN IL1RL.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96223957; PubMed=8626725;
 RA Parner P., Garika K.E., Bonnett T.P., Dower S.K., Sims J.E.;
 RT "IL-1RIP is a novel receptor-like molecule similar to the type I
 RT Interleukin-1 receptor and its homologues TI/ST2 and IL-1R AcP."
 RL J. Biol. Chem. 271:3967-3970(1996).
 DR EMBL: U43673; AAC52457.1; -.
 DR MGD: MGI:105383; IL1RL.
 DR InterPro: IPR003600; Ig_Like.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR004075; IntrIkn1_receptor.
 DR InterPro: IPR00157; TIR.
 DR Pfam: PF00047; Ig_3.
 DR Pfam: PF01582; TIR; 1.
 DR PRINTS: PR01537; INTRLKN1RLF.
 DR SMART: SM00410; IG_Like; 1.
 DR SMART: SM00255; TIR; 1.
 KW Signal.
 FT SIGNAL.
 SQ SEQUENCE 537 AA; 61600 MW; B119FEA03355458 CRC64;

Query Match

Best Local Similarity 27.9%; Score 254; DB 11; Length 537;
 Cons 95; Conservative 62; Mismatches 113; Indels 70; Gaps 18;

OY 9 PDFLSPEDOVLRPALGSSVALNCTA-----WVSGPHCSLPSVQ-----WLK 51
 DB 212 PALGPKCEKV-GVELGKQVNLCSALNKDLEFWYSIKRESDSDPNQEDRETTWLS 270
 OY 52 DGLPLGIGHYSLHEYSWKANLSEVLVSSVLCVNTSTEVYGAFTCSION---ISFSS 108
 DB 271 EG-----KLHASKILRFQ-----KITENYLANVLYNCTVANEBAIDTKSS 309
 OY 109 TLQR---AGPTSHVAVLASLVL--LALLLAALLVYKRLANVLYQD-AYGEVEING 162
 DB 310 VLVRKEIPIDPGHVFYGVYVLAASVAVCYITLCYIKVDLVLYRRIAREDEFTLIG 369
 OY 163 KLYDAVYS-SDC-PDRKRVNFILK--PO-LERRGYKFLDPRDLPRAPESADLVN 217
 DB 370 KTYDAVYSILKECPENKEVETLPRVLEKQFGYKLCIFERDVVPGAVVEIHSL 429
 OY 218 LSCRRLIYVLSAFLSRAMCSHSFREGJCR-LLELTRPIFTFGQRDPAN----- 270
 DB 430 IKSRLIYVLSQSYLTNG-ARRELSGLHEALVERKIKIILIEF-----TPASNTFLP 483
 OY 271 PALRLRQHRHLVTLIMRPGSVTPSSDFMKEVQALPRK 310
 DB 484 PSLKLKSTR---VLKWRADSPSANSRFWKNLVYIMPAK 519

RESULT 9

O13478 PRELIMINARY; PRT: 541 AA.
 AC 013478;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE IL-1RIP PRECURSOR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96223957; PubMed=8626725;
 RA Parner P., Garika K.E., Bonnett T.P., Dower S.K., Sims J.E.;
 RT "IL-1RIP is a novel receptor-like molecule similar to the type I
 RT Interleukin-1 receptor and its homologues TI/ST2 and IL-1R AcP."
 RL J. Biol. Chem. 271:3967-3970(1996).
 DR EMBL: U43672; AAC50390.1; -.
 DR InterPro: IPR003599; Ig_3.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR004075; IntrIkn1_receptor.
 DR InterPro: IPR000157; TIR.
 DR Pfam: PF00047; Ig_3.
 DR Pfam: PF01582; TIR; 1.
 DR PRINTS: PR01537; INTRLKN1RLF.
 DR SMART: SM00409; IG_1.
 DR SMART: SM00255; TIR; 1.
 KW Signal.
 FT SIGNAL.
 SQ SEQUENCE 541 AA; 62304 MW; 7173DB9C7EA71D32 CRC64;

Query Match

Best Local Similarity 28.0%; Score 231.5; DB 4; Length 541;
 Cons 99; Conservative 62; Mismatches 126; Indels 67; Gaps 19;

OY 6 DRA---PDFLSPEDOVLRPALGSSVALNCTA-----WVSGPHCSLPSVQWLKDG 53
 DB 209 DRSNIYVPLLGRKLNIV-AVELGKNVRLNCSALNEDVIYKMGFGENSDPRINHEKEM 267
 OY 54 LPLGIGHYSLHEYSWKANLSEVLVSSVLCVNTSTEVYGAFTCSIONISFSSFTL 110
 DB 268 RIMTPEG-----KMHASKVLRINEMIGESNLNV-----LYNCTVASTGTGDTKSFIL 313
 OY 111 QR---AGPTSHV--AAVLASLVLALLLAALLVYKRLANVLYQDAXGEYF-INDK 163
 DB 314 VRKADMDIPGHVFTRGMTIAVLIVAVVCLVYVYIVDVLFFYRHLTRDEFTLIGK 373
 OY 164 KYDAVYS-SDC---PEDRKRVNFILKPOLERRRGYKFLDPRDLPRAPESADLVN 218
 DB 374 TYDAVYSILKECPENKEVETLPRVLEKQFGYKLCIFERDVVPGAVVDEIHSL 433
 OY 219 SRCRLIYVLSAFLSRAMCSHSFREGJCR-----LLELTRPIF-ITFGQRDP 268
 DB 434 EKSRLIYVLSKYSYMSNEVRYELSGHEALVERKIKIILIEPT--PYNDFFFLPQ----- 487
 OY 269 AHPALMLRQHRHLVTLIMRPG-SVTPSSDFMKEVQALPRKVRYPVGPQ 321
 DB 488 ---SLKLKSHR---VLKWKADKSLSYNSRFWKNLVYIMPAKT-VKPGRDEPE 533
 RESULT 10
 O14915 PRELIMINARY; PRT: 570 AA.
 AC 014915;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE IL-1 RECEPTOR ACCESSORY PROTEIN (MEMBRANE INTERLEUKIN 1 RECEPTOR
 DE ACCESSORY PROTEIN).
 GN IL1RAP.
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
RN NCBI_TaxID=9606; (1)
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Huang J., Gao X., Li S., Cao Z.;
RN Proc. Natl. Acad. Sci. U.S.A. 0:0-0(1997).
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Saito T., Seki N.;
RN Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.
RN (3)
RP SEQUENCE OF 118-179 FROM N.A.
RC MEDLINE=98140136; PubMed=9479509;
RA "Dale M., Hammond D.W., Cox A., Nicklin M.J.H.;
RN "The human gene encoding the interleukin-1 receptor accessory protein
RT (ILRAK) maps to chromosome 3q28 by fluorescence in situ hybridization
RT and radiation hybrid mapping.";
RN Genomics 47:325-326(1998).
RP SEQUENCE FROM N.A.
RA Jensen L.P., Muzic D., Mantovani A., Whitehead A.S.;
RN "IL-1 signaling in the liver cells and the involvement of a
RT soluble form of the IL-1 receptor accessory protein.";
RL J. Immunol. 164:5277-5286(2000).
DR EMBL: AF029213; BAB84059.1; -
DR EMBL: AB006537; BAA25421.1; -
DR EMBL: AF016261; AAC39609.1; -
DR EMBL: AF167342; AAF71689.1; -
DR EMBL: AF167335; AAF71689.1; JOINED
DR EMBL: AF167337; AAF71689.1; JOINED
DR EMBL: AF167338; AAF71689.1; JOINED
DR EMBL: AF167339; AAF71689.1; JOINED
DR EMBL: AF167340; AAF71689.1; JOINED
DR EMBL: AF167341; AAF71689.1; JOINED
DR InterPro: IPRO03599; IG.
DR InterPro: IPRO03506; IG_MHC.
DR InterPro: IPRO04075; InterLknl_receptor1.
DR InterPro: IPRO00157; TIR.
DR Pfam: PF00047; IG_3.
DR Pfam: PF01582; TIR_1.
DR PRINTS: PRO1537; INTERLKNLIP.
DR SMART: SM00409; IG_2.
DR SMART: SM00255; TIR_1.
SW RECEPTOR.
SS SEQUENCE 570 AA; 65418 MW; 5F47F8D0CEA9B8A CRC64;
Query Match 10.6%; Score 228.5; DB 4; Length 570;
Best Local Similarity 27.6%; Pred. No. 4,1e-13;
Matches 91; Conservative 60; Mismatches 150; Indels 29; Gaps 11
OY 9 PDLIPSEDDOYLRLPALGSSVALNCTAWVVSFGHCLPSYOWLJKG-LPLIGIGHYSHEV 67
DB 243 FVHSIPNDHVVYERKEPEELLICFYV-FSLPMDSRNFWMTIGKRPDDITDVTINE- 300
OY 68 SWKANIISEVLSVSLGV-NVSTLEYGAFTC--SIQNISSFSSFTLORAGPTSHVAAYV 123
DB 301 SISRSREDETRQIILSIKRVTSIEDIKRKYVCHARSAGEVAKAKOKVPAARYTEL 360
OY 124 A--SLVLLALLLAALLLYKRCRLNVLLMYODAYGEV-INDGKLYDAVYSSDCEDRK 179
DB 361 ACGGATVLLVLLIVVYHYV-LEMLVLYRNHGTDETDINDGKEYIIVYSVRNBEER 419
OY 180 FVNFTIKPOLERRRGKFLFLDRDILLPRAEDSADLLVNLISCRLLIYVLSDAELSRAMCS 239
DB 420 FVLLITRGVLENEFGKLCIFDRDSIPGIVYDETLSFIQSRRLVLSPNVY----- 473
OY 240 HSRREGKRLLETTRKPTITEG-----QRDPAPRLRLIQHRLVLTLLRL-PG 291

Dd	474	-	LOGTQALLETLKAGLEENMASRCGINIVILVYKAKVEIKYRKELRAKIAVLTVIWMKEG	530
Oy	292	SVPSSDPEKWEVOLAPRKRYRVEDDPQ	321	
		: :		
Dd	531	SKYPQGRFWKLOLVAMPVKRSPPRSSDDQ	560	
 RESULT 11				
ID	Q9ERS7	PRELIMINARY:	PRT:	574 AA.
AC	Q9ERS7;			
DT	01-MAR-2001	(TREMBLrel. 16,	Created)	
DT	01-MAR-2001	(TREMBLrel. 16,	Last sequence update)	
DT	01-DEC-2001	(TREMBLrel. 19,	Last annotation update)	
DE	IL-IRRP2.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
NCBI_TaxID	10090.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20459050; PubMed=10882729;			
KA	Born T.L., Smith D.E., Garka K.E., Renshaw B.R., Bertles J.S.,			
RA	Shim J.E.,			
KT	Identification and characterization of two members of a novel class			
KT	of the Interleukin-1 receptor (IL-1R) family. Delineation of a new			
KL	U.S. Bio. Chem. 273:29946-29954(2000).			
DR	HSPB, AF244433; AAG21367.1; "			
DR	HSPB, F14778.1;IRA.			
DR	InterPro: IPR003399; Ig_C2.			
DR	InterPro: IPR003396; Ig_C2.			
DR	InterPro: IPR003600; Ig_Like.			
DR	InterPro: IPR000157; Ig_MHC.			
DR	Pfam: PF00582; TIR_2.			
DR	SMART: SMO0409; IG_3.			
DR	SMART: SMO0408; IG_C2_2.			
DR	SMART: SMO0410; IG_Like; 1.			
DR	SMART: SMO0255; TIR; 1.			
KW	Immunoglobulin domain.			
SEQ	SEQUENCE 574 AA: 65108 MW; A677A77BBFA50A76 CRC64;			
 Query Match 10.1%; Score 217; DB 11; Length 574; Best Local Similarity 23.3%; Pred. No. 5,3e-12; Matches 100; Conservative 62; Mismatches 15; Indels 132; Gaps 20.				
Oy	13	SPSEDQYLRIALCSVAALNCTAMVVSQGPICSPSVQWLK---	DGLEP----	56
		: :		
Dd	131	SPDYOOILP-IKGSGSLNCHLYFPES--CALDSIKWKGEELKACKTKYSPGAILVN	187	
Oy	57	-----GIGGHYSLHEY-----SWKANLSSEGLVSYV	82	
		: :		
Dd	188	NVAVEDGSYACSARLTILGRHFTIRNYIAVNTKEVEYGRRIPIITYPKNNSTLEVLLG	247	
Oy	83	LGV--NVTSIEVYGAFTC-----SIO-----NISFS	107	
		: :		
Dd	248	LIVNCNTDIETKENTILCRKWVNNLTVDVYKDSKRIOEGEITFNSLRDQIRTYVTIFLK	307	
Oy	108	FLLORAGP--TSNV-----AAVLASLVLALLAL--LYKCALTAVL	147	
		: :		
Dd	308	VKMEDYGRPTCHAGVASAYILLIYPVDFPAFYILGIMLFELLVVSFLPIYSFKDIM	367	
Oy	148	LWQODVYGEVEI--NGDKILDVAVSYSDCPEDRK-----RVNFILKPOLERRRGYLLFD	200	
		: :		
Dd	368	LWRSAFHNAQAFDEKIIDDAVVLPRKYRQSQHGDVDTVLILKEPVLEKOGGYLPLF	427	
Oy	201	DRDLLEPAEPSADLLVNISRCRRILTVLSDAFLSRACSHSPREGC---RLLELTRRPI	257	
		: :		
Dd	428	GREFPGQAVASYTDENIKLCRRILMVFAEPSSFGFLKNLSEEOIAYNALTOHGMMKI	487	
Oy	258	FTTECGQRDPAPH--ALRIIRQHRILVTLILMRGVSVTPSSD-----FWKEYOLALPKV	311	

01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE INTERLEUKIN-1 RECEPTOR ACCESSORY PROTEIN.
 GN IL-1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9626535; PubMed=8964912;
 RA Liu C., Chalmers D., Maki R., De Souza E.B.:
 RT "Rat homolog of mouse interleukin-1 receptor accessory protein:
 cloning, localization and modulation studies.";
 RL J. Neuroimmunol. 66:41-48(1996).
 DR EMBL: U48592; AAB03502.1; -;
 DR InterPro: IPR003599; IG_1;
 DR InterPro: IPR004075; IG_MHC.
 DR InterPro: IPR000157; TIR.
 DR Pfam: PF00047; IG_3.
 DR Pfam: PF01582; TIR; 1.
 DR Pfam: PF01537; INTRILK1R1F.
 DR SMART: SM00409; IG; 2.
 DR SMART: SM00255; TIR; 1.
 KW Receptor.
 SO SEQUENCE 570 AA; 65598 MW; 20C3A5478127AABE CRC64;

Query Match 9.8%; Score 210.5; DB 11; Length 570;
 Best Local Similarity 25.5%; Pred. No. 2.2e-11;
 Matches 86; Conservative 64; Mismatches 144; Indels 43; Gaps 13;
 OY 9 PDLFSSDQVLRPALGSSVALNCTA--VYSGPHCSLPYOMLKDG-----PLIGIGH 61
 DB 243 PIYSPNDRVYVEKPEBELVIPCXYVFSIMDSH---NEIWIIDCKKDDVDVITII 299
 OY 62 YSLHYSWKANKISLVLSVGLV-NVTSPEYGAFTCSION----ISFSFTLORAGP- 115
 DB 300 ESV-SYS---STDEFTQTOLSTIKKVTPELKRNYCHARNAGEAEQAKKOKATIP 354
 OY 116 --TSVAVAVLSLVALLAALLVYKCRNLVLYODAYGEV-INDKRLDAYSYS 172
 DB 355 RYVLEACGFGATVFLVYVLI--VYHYVLEMYLFTFAHFGIDETLLDGKEDIVSYA 412
 OY 173 DCPERKRVNFTLKPOLERRRGYKFLDDRLPRAPSADLVNLSRCRLIVYSDAF 232
 DB 413 RNAEEEFVLLTLKGVLENEFGYKLCIFDRSPFGIVTDETLISFICKSRRLVLSPNY 472
 OY 233 LSRWCSHSPREGLCRLLELTRRPIFTFEG-----QRDPAPALRLRLROHRLVTL 285
 DB 473 V-----LQGTQALLELAKLENNASRGINIVILVQYKAVKDLKVKELKRAKSVLT 523
 OY 286 LLMR-PGSVTPSSSDFWKEVOLALPRKYVYRVEGDPQ 321
 DB 524 IKWGEKSKYPOGRFWKOLQYAMPVKKSPWSSSDKQ 560

RESULT 15
 O9DEES PRELIMINARY; PRT; 533 AA.
 AC O9DEES;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 ST 2L.
 GN ST2.
 OS Gallus gallus (Chicken).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA "Iwahana=H.;

*Chicken str2 gene: genomic structure and its products.";
 RT Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AB041738; BAB20773.1; -;
 DR EMBL: IPR003599; IG_1;
 DR InterPro: IPR003599; IG_1;
 DR InterPro: IPR003598; IG_C2.
 DR InterPro: IPR003600; IG_Like.
 DR InterPro: IPR003006; IG_MHC.
 DR InterPro: IPR000157; TIR.
 DR Pfam: PF00047; IG_2;
 DR Pfam: PF01582; TIR; 1.
 DR SMART: SM00409; IG; 3.
 DR SMART: SM00408; IG_C2; 2.
 DR SMART: SM00410; IG_Like; 1.
 DR SMART: SM00255; TIR; 1.
 KW Immunoglobulin domain.
 SO SEQUENCE 533 AA; 61755 MW; AB068E5A763EFCE4 CRC64;

Query Match 9.8%; Score 210; DB 13; Length 533;
 Best Local Similarity 23.5%; Pred. No. 2.2e-11;
 Matches 76; Conservative 75; Mismatches 142; Indels 30; Gaps 13;
 OY 12 LSPSEDOVLRPALGSSVALNCTA--VYSGPHCSLPYOMLKDGPLIGIGHYSL-HEYS 68
 DB 211 LPPKDEVDVLEVLGAFAFSIKCOARLGIKKOP---IAVYTDVNMKPAVYADEEKIROEFS 267
 OY 69 WVKANISEVLSVGLV-NVTSPEYGAFTCSIONISFSFTL-----ORAGPTSHVA 121
 DB 268 YPEGQHOEYEGATLTISNIETDLOSFSVCVAMNEMGNTRFTVTLRLKKSGP--NVLM 325
 OY 122 VLASLVALLAALLVYKCRNLVLYODAYGEV-INDKRLDAYVSY-----SDCP 175
 DB 326 IAGFLVFLMSVAVSVVLSQSPFVDVLLYRLDFQAYSKDCKITDYAVVYPRSHTEAT 385
 OY 176 EDKRVNFTLKPOLERRRGYKFLDDRLPRAPSADLVNLSRCRLIVYSDAFSLR 235
 DB 386 FVEYVYQIMPDILNKCGYTLCTIGRTYDGEDKANAIEKRSKRLIILIQOLINC 445
 OY 236 ANCSHSPREGLCRLLELTRRPIFTFEGQ--RRDPAPALRLRLROHRLVTLIMR-PGS 292
 DB 446 QELNIDQIATLALNLDNDTKVILEMEKMSYENLOESLRYTIKQGGYK--WKQHT 502
 OY 293 VTP---SDFWKEVOLALPRKYR 312
 DB 503 VHPQSSNSKFMKHYRHPVTLR 525

Search completed: November 7, 2002, 18:51:50
 Job time : 74 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 7, 2002, 18:41:51 ; Search time 26 Seconds
(without alignments)
610.578 Million cell updates/sec

Title: US-09-598-443-2
Perfect score: 2147
Sequence: 1 MGVCDRAPDLSPEDQVL.....GSRNVSARTDFYCLVSKDDM 410

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Sead: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Swissprot_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	209.5	9.8	569	IL1R_HUMAN	P14778 homo sapien
2	190.5	8.9	576	IL1R_MOUSE	P13504 mus musculus
3	176	8.2	1097	TOLL_DROME	P08953 drosophila
4	168.5	7.8	576	IL1R_RAT	O02955 rattus norv
5	152	7.1	841	IL1R_BOVIN	O96165 bos taurus
6	144	6.7	841	IL1R_MOUSE	O96165 mus musculus
7	142.5	6.6	826	IL1R_PAPAN	O95822 papio anubi
8	142	6.6	833	IL1R_FELCA	P58727 felis silve
9	138	6.4	839	IL1R_HUMAN	O00206 homo sapien
10	135.5	6.3	839	IL1R_BOVIN	O95149 bos taurus
11	134	6.2	839	IL1R_PAPAN	O95149 papio anubi
12	131	6.1	781	IL1R_CHICK	O96165 gallus gall
13	131	6.1	784	IL1R_MACRA	O95853 macaca fasc
14	131	6.1	793	IL1R_CHICK	O96165 gallus gall
15	129	6.0	843	IL1R_HORSE	O96165 equus caball
16	125	5.8	784	IL1R_HUMAN	O00206 homo sapien
17	124.5	5.8	808	IL1R_MOUSE	O03142 mus musculus
18	121	5.6	784	IL1R_CRICR	O96165 cricetus
19	117	5.4	795	IL1R_MOUSE	O96165 mus musculus
20	116	5.4	835	IL1R_RAT	O96165 rattus norv
21	116	5.4	838	IL1R_CRICR	O96165 cricetus
22	115	5.3	296	MY88_MOUSE	P22366 mus musculus
23	114.5	5.3	786	IL1R_HUMAN	O15399 homo sapien
24	114.5	5.1	795	MY88_HUMAN	O96165 homo sapien
25	110.5	5.1	795	IL1R_MOUSE	O96165 mus musculus
26	110.5	5.1	835	IL1R_MOUSE	O96165 mus musculus
27	109.5	5.1	1032	IL1R_HUMAN	O96165 homo sapien
28	107.5	5.0	1032	IL1R_HUMAN	O96165 homo sapien
29	106.5	5.0	859	IL1R_HUMAN	O96165 homo sapien
30	103.5	4.8	858	IL1R_MOUSE	O60602 mus musculus
31	103	4.8	365	CXAR_HUMAN	P78310 homo sapien
32	102.5	4.8	802	FCR4_HUMAN	P22455 homo sapien
33	100.5	4.7	629	YS50_MYCTU	O05809 mycobacteri

34	100	4.7	1032	1	TLR8_MOUSE	P58682 mus musculus
35	99	4.6	811	1	TLR4_HUMAN	O96165 mus musculus
36	98	4.6	1032	1	TLR9_MOUSE	O96165 mus musculus
37	98	4.6	1041	1	TLR8_HUMAN	O96165 mus musculus
38	96.5	4.5	766	1	SM7_BRARE	O96165 mus musculus
39	96	4.5	288	1	PH2_MOUSE	O15296 mus musculus
40	95	4.4	676	1	LX1B_HUMAN	O15296 mus musculus
41	94	4.4	2012	1	DSCA_HUMAN	O60469 mus musculus
42	93.5	4.4	1709	1	SN_HUMAN	O96165 mus musculus
43	93	4.3	764	1	ICCR_DROME	O08180 drosophila
44	92	4.3	729	1	AD21_MOUSE	O96165 mus musculus
45	92	4.3	1694	1	SN_MOUSE	O62230 mus musculus

ALIGNMENTS

RESULT 1
ID IL1R_HUMAN
AC P14778:
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Interleukin-1 receptor, type I precursor (IL-1R-1) (IL-1R-alpha)
DE (P80) (Antigen CD121a).
GN IL1R1 OR IL1RA OR IL1R.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBT_Taxid-9606;
[1]
RP TISSUE=Liver;
RC MEDLINE=90098789; PubMed=2532321;
RA Chua A.O., Gubler U.;
RT "Sequence of the cDNA for the human fibroblast type Interleukin-1
RT receptor".
RL Nucleic Acids Res. 17:10114-10114(1989).
[2]
RN SEQUENCE FROM N.A.
RP TISSUE=T-cell;
RC MEDLINE=90046906; PubMed=2530587;
RA Sims J.E., Acres R.B., Gubler U.;
RT "Cloning the Interleukin 1 receptor from human T cells".
RL Proc. Natl. Acad. Sci. U.S.A. 86:8946-8950(1989).
[3]
RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 18-332 IN COMPLEX WITH IL1B.
RP MEDLINE=97215903; PubMed=9062193;
RA Vigers G.P., Anderson L.J., Cafes P., Brandhuber B.J.;
RT "Crystal structure of the type-I Interleukin-1 receptor complexed
RT with Interleukin-1beta".
RL Nature 386:190-194(1997).
[4]
RN X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 21-331 IN COMPLEX WITH IL1RA.
RP MEDLINE=97215904; PubMed=9062194;
RA Schneider H., Tardif C., Trump-Kallmeyer S., Soffientini A.,
RA Sarubbi E., Akeson A., Bowlin T., Yanofsky S., Barrett R.W.;
RT "A new cytokine-receptor binding mode revealed by the crystal
RT structure of the IL-1 receptor with an antagonist".
RL Nature 386:194-200(1997).
CC -!- FUNCTION: RECEPTOR FOR INTERLEUKIN-1 ALPHA (IL-1A), BETA (IL-1B),
CC AND INTERLEUKIN-1 RECEPTOR ANTAGONIST PROTEIN (IL-1RA). BINDING TO
CC THE AGONIST LEADS TO THE ACTIVATION OF NF-KAPPA B.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -!- DATABASE: NAME=PROV: NOTE=CD guide CD121a entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd121a.htm".
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
CC or send an email to license@ebi.ac.uk).

DR	EMBL; X16896; CAA34773.1; -
DR	EMBL; M27492; AAS59137.1; -
DR	PIR; S06928; S06928.
DR	PIR; A36187; A36187.
DR	PDB; 1IRB; 04-FEB-98.
DR	PDB; 1IRA; 17-JUN-98.
DR	MIM; 147810; -
DR	InterPro; IPR003006; IG_MHC.
DR	InterPro; IPR003600; IG_1ike.
DR	InterPro; IPR004076; IGIL1_receptor.Pre.
DR	InterPro; IPR004075; IGIL1_receptor.
DR	InterPro; IPR004074; InterleukinReceptor1L1.
DR	InterPro; IPR00157; TIR.
DR	Pfam; PF00047; Ig_2.
DR	Pfam; PF01582; TIR; 1.
DR	PRINTS; PRO1538; INTERLEUKIN1L1.
DR	PRINTS; PRO1536; INTERLEUKIN1ZF.
DR	PRINTS; PRO1537; INTERLEUKIN1F.
DR	SMART; SM00410; IG_1like; 1.
DR	SMART; SM00255; TIR; 1.
DR	PROSITE; PS50104; TIR; 1.
KW	Immunoglobulin domain; Transmembrane; Glycoprotein; Receptor; Signal. Repeat; 3D-structure.
KW	Repeat; 3D-structure.
FT	CHAIN 1 17
FT	SIGNAL 18 569
FT	DOMAIN 18 336
FT	TRANSMEM 337 356
FT	DOMAIN 357 569
FT	DOMAIN 37 103
FT	DOMAIN 135 203
FT	DOMAIN 241 319
FT	DOMAIN 383 541
FT	DISULEID 23 104
FT	DISULEID 44 96
FT	DISULEID 121 164
FT	DISULEID 142 196
FT	DISULEID 248 312
FT	CARBOHYD 130 100
FT	CARBOHYD 133 193
FT	CARBOHYD 229 233
FT	CARBOHYD 249 249
FT	CARBOHYD 253 263
FT	CARBOHYD 297 297
SO	SOURCE 569 AA; 65402 MW; 5BAAB3FE0225C25 CRC64;

[illegible]

Db 439 VINENYKSRLLIIVRTSSEFSSGSSSEQYAMNMYDQGIKVYLLLELEKIQDEK 496

Qy 255 RPIFIPEQGRDPAHPALRLRHRILVLLM---RGSVITSDDKWEVOLAP 308

Db 499 MPESIKFIKQK---HGAIK-----WSGSDTQGSQSAIKRQKVRIRNI 539

AC	IL1R_MOUSE	STANDARD:	PRt:	576 AA.
AD	P13504;			
DT	01-JAN-1990 (Rel. 13, Created)			
DT	01-JAN-1990 (Rel. 13, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Interleukin-1 receptor, type I precursor (IL-1R-1) (P80).			
GN	IL1R1 OR IL1RA OR IL-1RL.			
OS	Mus musculus (mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 20-45.			
RX	MEDLINE=88290679; PubMed=2969618;			
RA	Sims J.E., March C.J., Cosman D., Widmer M.B., McDonald H.R.,			
RA	McLachan C.J., Grubin C.E., Wignall J.M., Jackson J.L., Call S.M.,			
RT	Friend D., Albert A.R., Gillis S., Urdal D.L., Dower S.K.,			
RT	"cDNA expression cloning of the IL-1 receptor, a member of the			
RL	immunoglobulin superfamily.";			
RL	Science 241:585-589(1988).			
RP	[2]			
RP	PHOSPHORYLATION AT THR-556			
RX	MEDLINE=91254338; PubMed=1628344;			
RA	Bird T.A., Woodward A., Jackson J.L., Dower S.K., Sims J.E.;			
RT	"Photol ester induces phosphorylation of the 80 kilodalton murine			
RT	Interleukin 1 receptor at a single threonine residue.";			
RL	Biochem. Biophys. Res. Commun. 177:61-67(1991).			
CC	-1- FUNCTION: RECEPTOR FOR INTERLEUKIN-1 ALPHA (IL-1A), BETA (IL-1B),			
CC	AND INTERLEUKIN-1 RECEPTOR ANTAGONIST PROTEIN (IL-1RA). BINDING TO			
CC	THE AGONIST LEADS TO THE ACTIVATION OF NF-KAPPA B.			
CC	-1- SUBCELLULAR LOCATION: Type 1 membrane protein.			
CC	-1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.			
CC	-1- SIMILARITY: CONTAINS 1 TIR DOMAIN.			
CC				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way that			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib-sib.ch).			
CC				
CC	EMBL: M20658; AAA9279.1; ..			
DR	PIR: A32604; A32604.			
DR	HSP: P14778; IL1R.			
DR	MGD: MGI:96545; IL1rl.			
DR	Interpro: IPR003006; I9_MHC.			
DR	Interpro: IPR003600; I9_1like.			
DR	Interpro: IPR004076; Inlkl1_receptor1_pre.			
DR	Interpro: IPR004075; Inlrlkl1_receptor1.			
DR	Interpro: IPR004074; Inlrlkl1_receptor1_1l.			
DR	Interpro: IPR000157; TIR.			
DR	Pfam: PF00047; I9; 2.			
DR	Pfam: PF01582; TIR; 1.			
DR	PRINTS: PR01538; INTERLEUKN1R1.			
DR	PRINTS: PR01536; INTERLEUKN1R2F.			
DR	PRINTS: PR01537; INTERLEUKN1R1F.			
DR	SMART: SM00410; IG_1like; 1.			
DR	SMART: SM00255; TIR; 1.			
DR	PROSITE: PS20104; TIR; 1.			
DR	Immunoglobulin domain; Receptor; Glycoprotein; Transmembrane; Signal;			
KW	Repeat; Phosphorylation.			
FT	SIGNAL			
FT	CHAIN 20 576 INTERLEUKIN-1 RECEPTOR, TYPE I.			

```

FT DOMAIN 20 338 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 339 359 POTENTIAL.
FT DOMAIN 360 576 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 39 113 IG-LIKE C2-TYPE DOMAIN 1.
FT DOMAIN 138 206 IG-LIKE C2-TYPE DOMAIN 2.
FT DOMAIN 244 322 IG-LIKE C2-TYPE DOMAIN 3.
FT DOMAIN 386 344 TIR.
FT DISULFID 25 107 BY SIMILARITY.
FT DISULFID 46 99 BY SIMILARITY.
FT DISULFID 145 199 BY SIMILARITY.
FT DISULFID 251 315 BY SIMILARITY.
FT CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 236 236 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 266 266 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 PHOSPHORYLATION (BY PC).
SC SEQUENCE 576 AA: 66697 MW: 7AA8304C86412A16 CRC64:

```

```

Query Match 8.9% Score 190.5; DB 1: Length 576;
Best Local Similarity 24.0%; Pred. No. 1,1e-08;
Matches 86; Conservative 71; Mismatches 151; Indels 51; Gaps 18;

```

```

OY 6 DRAPDFLSPEDQVLRPALGSSVALNCTAMVVSQPHCSLPYQWLMKGLPLGIGHYSLH 65
DB 227 DR-PVLSP-RMNETIADGSMIQLCN---VTGQFSDL-VYMKNGSEIENMDFLAE 279
OY 66 EYSWAK--ANLSEVLYSVLVGNVSTEVYGAFTCSIONIS-FSSFTLQRAQPTSHVAA 121
DB 280 DVQFVHPSTKRYTLITLTLNISEVKSQFYRPFICVVKNTIFESAHVOLITVPDFKN 339
OY 122 VLASLLVL--ALLAALLVYKRLNVLWYODA---YGEVEINDKRIYDAVY----- 169
DB 340 YLIGGITLITATIVCCVCYIKFKKVDIVMYRSCGFLPSKASDSKTYDAYLYPKTIG 399
OY 170 --SYSDPEDRKVFNLIPOLERRRGYKFLDRLDLPRAEPSADLLVLSRCRLIYV 227
DB 400 EGSFOLD--TFVFKLLEPVELEGQGYKFLFYGRDGYGEDEVTEVNEVKKSRLLII 456
OY 228 LSAFLSRMCSISFREGICRLLELTRRPIFTT---EGQRDPHAPALRLRLRHRILV 283
DB 457 LVNDMGFMLOQSSSEQAIYNALIOEGIKITYLELEKTIODEKPKADISIOFTKQ-KHGV 515
OY 284 TLLTW-----PGSVTPSSDFMKVEYQALLPRKVR-----YRPVEDPQTQLODDKDPM 331
DB 516 --ICWSGDFQERPOSA--KTRFMKRLRYQMPAPORRSPLSKHLRLTLDP--VADTKRKL 567

```

RESULT 3

```

TOLL_DROME STANDARD; PRT: 1097 AA.
AC P08953; Q9YBB8;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE TOLL protein precursor.
GN TL OR CS5490.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
ON NCBI_TaxID=7227;
RX MEDLINE=8135760; PubMed=2449285;
RA Hashimoto C., Hudson K.L., Anderson K.V.;
RT "The Toll gene of Drosophila, required for dorsal-ventral embryonic
RT polarity, appears to encode a transmembrane protein."
RL Cell 52:269-279(1988).
RN [2]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt J.G., Nelson C.R., Milos G.L.G.,
RA Abril J.F., Baybayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borrova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos J.M., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dong L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasro P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Maltel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mlshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
RA Palazolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svrtka R., Tector C., Turner R., Venter G.M., Weissbach J.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RL "The genome sequence of Drosophila melanogaster.";
RN Science 287:2185-2195(2000).
RP FUNCTION.
RX MEDLINE=91092252; PubMed=2124970;
RA Keith F.J., Gay N.J.;
RT "The Drosophila membrane receptor Toll can function to promote
RT cellular adhesion."
RL EMBO J. 9:4299-4306(1990).
CC -!- FUNCTION: REQUIRED FOR DORSAL-VENTRAL EMBRYONIC POLARITY. MAY
CC -!- FUNCTION AS A MEMBRANE RECEPTOR. PROMOTES HETEROPHILIC CELLULAR
CC ADHESION.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: CONTAINS 21 LEUCINE-RICH REPEATS (LRR).
CC -!- SIMILARITY: CONTAINS 1 TIR DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
DR EMBL: M19669; AAA28941.1; -
DR EMBL: AE003758; AAF56624.1; -
DR PIR: A28943; A28943.
DR Flybase: FBgn0003717; T1.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000483; LRR_Cterm.
DR InterPro: IPR000372; LRR_Nterm.
DR InterPro: IPR003592; LRR_out.
DR InterPro: IPR003591; LRR_Typ.
DR InterPro: IPR000157; TIR.
DR Pfam: PF00560; LRR; 11.

```

[illegible]

Oy	284	TL1MRGSGTSPSSD--FMKEVQLAKPKRYAVRPEDDPTQLOADKDPMLIRLGRVPGC	341
Db	976	TYLKW-----GDWFWDKLRFLALPHR---RV-----GNIGNG	1005
Oy	342	RALDSEVDPDPEDDLGVRGPFGESSAPHTS	373
Db	1006	ALITAIKSGSTDKLELIRK---SPVTPPLTIT	1034
RESULT 4			
TL1R-RAT	TL1R-RAT	STANDARD:	PRT: 576 AA.
AC	002953;		
DT	01-JUL-1993	(Rel. 26, Created)	
DT	01-JUL-1993	(Rel. 26, Last sequence update)	
DT	01-MAR-2002	(Rel. 41, Last annotation update)	
DE	Interleukin-1 receptor, type I precursor (IL-1R-1) (P80).		
GN	IL1R1 OR IL1RA.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
OX	NCBI_TaxID=10116;		
RN	11		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=ganglion;		
RA	MEDLINE=9326794; PubMed=7684399;		
RA	Hart R.P., Liu C., Shadick A.M., McCormack R.J., Jonakal G.M.;		
RT	an mRNA homologous to interleukin-1 receptor type I is expressed in		
RT	cultured rat sympathetic ganglia."		
RL	J. Neuroimmunol. 44:49-56(1993).		
CC	-1- FUNCTION: RECEPTOR FOR INTERLEUKIN-1 ALPHA (IL-1A), BETA (IL-1B) TO		
CC	AND INTERLEUKIN-1 RECEPTOR ANTAGONIST PROTEIN (IL-1RA). BINDING TO		
CC	THE AGONIST LEADS TO THE ACTIVATION OF NF-KAPPA B.		
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein		
CC	-1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.		
CC	-1- SIMILARITY: CONTAINS 1 TIR DOMAIN.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@sib-sib.ch).		
CC	-----		
DR	EMBL; M95578; AAL6196.1; ALT_SEQ.		
DR	HSSP; P14778; IIRA.		
DR	Interpro; IPR003006; I9_MHC.		
DR	Interpro; IPR003600; I9_1like.		
DR	Interpro; IPR004076; IRLK1_receptorI_pre.		
DR	Interpro; IPR004075; IRLK1L_receptorI.		
DR	Interpro; IPR004074; IRLK1L_receptorI_II.		
DR	Interpro; IPR000157; TIR.		
DR	Pfam; PF00047; I9; 2.		
DR	Pfam; PF01582; TIR; 1.		
DR	PRINTS; PRO1538; INTERLEUKN1R1.		
DR	PRINTS; PRO1536; INTERLEUK1R12F.		
DR	PRINTS; PRO1537; INTERLEUK1R12F.		
DR	SMART; SM00410; IG_1like; 1.		
DR	SMART; SM00255; TIR; 1.		
DR	PROSITE; PS50104; TIR; 1.		
KW	Immunoglobulin domain; Receptor; Glycoprotein; Transmembrane; Signal;		
KW	Repeat; Phosphorylation.		
FT	CHAIN	1..19	
FT	FT	20..576	BY SIMILARITY.
FT	DOMAIN	20..338	INTERLEUKIN-1 RECEPTOR, TYPE I.
FT	TRANSMEM	339..359	EXTRACELLULAR (POTENTIAL).
FT	DOMAIN	360..576	POTENTIAL.
FT	DOMAIN	39..113	CYTOSOLSMIC (POTENTIAL).
FT	DOMAIN	138..206	IG-LIKE C2-TYPE DOMAIN 1.
FT	DOMAIN	244..322	IG-LIKE C2-TYPE DOMAIN 2.
FT	DOMAIN	386..544	IG-LIKE C2-TYPE DOMAIN 3.
FT	TT	25..107	TIR.
FT	DISULFID		BY SIMILARITY.

```
FT DISULFID 46 99 BY SIMILARITY.
FT DISULFID 145 199 BY SIMILARITY.
FT DISULFID 251 315 BY SIMILARITY.
FT CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 236 236 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 266 266 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT MOD_RES 556 556 PHOSPHORYLATION (BY PRC) (BY SIMILARITY).
SQ SEQUENCE 576 AA: 66758 MW: 558E20C92385A34A CRC64;

Query Match 7.8%; Score 168.5; DB 1; Length 576;
Best Local Similarity 24.8%; Pred. No. 8.9e-07;
Matches 88; Conservative 56; Mismatches 128; Indels 83; Gaps 21;

QY 6 DRAPDPLSPEDQVLRPALGSSVALNCTAMVYSGPHCSLPYQWLKDGPLIGIGHYSLSH 65
DB 227 DR-PVIMSP-RNETMEADPGSTQLICN--VTGQPTDL--YVWKWNGSEIEMDDPILAE 279
QY 66 EYSWWK--ANLSEVLVSVLGAVNTSTEVYG-AFTCSIONISFSSFTLRAGPTSHVA 121
DB 280 DVGFLHPAKKRYTLIT-LNVSEVKSQFYRPFICFKNTHILE-----TAHVRL 330
QY 122 VLA-----SLVLLALLALLLYKC-----RLNVLVYQDAYGEV---EINDKLYDA 167
DB 331 VYVPDPEKRYLLIGGFAIFATAVFCACIKVFKVDIVLWYRDCSDPLPRKASDGRTPYA 390
QY 168 VYSYSCPEDRKPFV--NFLK--PQ-LERRRGYKFLDODRDLRAPEFSADLLVLSNC 221
DB 391 VYLPKTYEGSGAVYIDTFEYKFLPEVLEGQFYKFLICGRDVGEDTLEVNTENVKKS 450
QY 222 RLIVLVSDAFLSRAMCSHS-----FREG-LCHLEL-----TRRPIFTFEG 263
DB 451 RLIIITLVADMGSFSCLOGSSEGOIAIYALIREGKIITLLELEKIQDYEKMPESIOFLK 510
QY 264 QRRDPAHPLRLRQHRHLVTLILW-----RPGSVTPSSDEKVEVOLAPRKVR 312
DB 511 QK---HGA-----ICWSGDKERPOSA--KTRFKNLRYQMPAQR 546

RESULT 5
TIR4_BOVIN STANDARD; PRT; 841 AA.
ID TIR4_BOVIN
AC Q9GL65;
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DB 01-MAR-2002 (Rel. 41, Last annotation update)
E TIR4-like receptor 4 precursor.
GN TIR4.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN NP_059558.1;
RP SEQUENCE FROM N.A.
RA Guionaud C.T., Dubey C., Jungi T.W.;
RT "Bovine Toll-like receptor 4 (TIR4)".
RL Submitted (NOV-2001) to the EMBL/Genbank/DBJ databases.
CC -I- FUNCTION: Cooperates with MD-2 and CD14 to mediate the innate
CC immune response to bacterial lipopolysaccharide (LPS). Acts via
CC MyD88, TIRAP and TRAF6, leading to NF-kappa-B activation, cytokine
CC secretion and the inflammatory response. (By similarity).
CC -I- SUBUNIT: Belongs to the lipopolysaccharide (LPS) receptor, a
CC multi-protein complex containing at least CD14, MD-2 and TIR4.
CC Blinds MD-2 via the extracellular domain. Blinds MyD88 and TIRAP via
CC their respective TIR domains (By similarity).
CC -I- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -I- SIMILARITY: BELONGS TO THE TOLL-LIKE RECEPTOR FAMILY.
CC -I- SIMILARITY: CONTAINS 1 TIR DOMAIN.
CC -I- SIMILARITY: CONTAINS 18 LEUCINE-RICH REPEATS (LRR).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL: AF310952; AAG32061.2; -.
CC DR Interpro: IPR001611; LRR.
CC DR Interpro: IPR000483; LRR_Cterm.
CC DR Interpro: IPR003582; LRR_Out.
CC DR Interpro: IPR003581; LRR_Typ.
CC DR Interpro: IPR001571; TIR.
CC DR Pfam: PF00560; LRR; 3.
CC DR Pfam: PF01463; LRRCT; 1.
CC DR Pfam: PF01582; TIR; 1.
CC DR PRINTS: PRO0019; LEURICRPT.
CC DR SMART: SM00370; LRR; 1.
CC DR SMART: SM00082; LRRCT; 1.
CC DR SMART: SM00369; LRR_Typ; 3.
CC DR SMART: SM00255; TIR; 1.
CC DR PROSITE: PSS0104; TIR; 1.
CC KW Receptor; Immune response; Inflammatory response; Signal;
CC KW Transmembrane; Repeat; Leucine-rich repeat; Glycoprotein.
CC FT SIGNAL 1 23
CC FT CHAIN 24 841
CC FT DOMAIN 24 632
CC FT TRANSMEM 633 653
CC FT DOMAIN 654 841
CC FT REPEAT 53 76
CC FT REPEAT 77 100
CC FT REPEAT 102 124
CC FT REPEAT 149 173
CC FT REPEAT 174 197
CC FT REPEAT 203 225
CC FT REPEAT 277 300
CC FT REPEAT 310 334
CC FT REPEAT 350 372
CC FT REPEAT 373 398
CC FT REPEAT 400 421
CC FT REPEAT 422 445
CC FT REPEAT 446 469
CC FT REPEAT 471 494
CC FT REPEAT 495 518
CC FT REPEAT 520 542
CC FT REPEAT 544 566
CC FT REPEAT 568 592
CC FT DOMAIN 673 819
CC FT CARBOHYD 35 35
CC FT CARBOHYD 73 73
CC FT CARBOHYD 205 205
CC FT CARBOHYD 238 238
CC FT CARBOHYD 282 282
CC FT CARBOHYD 309 309
CC FT CARBOHYD 497 497
CC FT CARBOHYD 526 526
CC FT CARBOHYD 575 575
CC FT CARBOHYD 625 625
CC SQ SEQUENCE 841 AA: 96026 MW: C5E17CB9C798CD16 CRC64;

Query Match 7.1%; Score 152; DB 1; Length 841;
Best Local Similarity 24.4%; Pred. No. 3.9e-05;
Matches 87; Conservative 50; Mismatches 118; Indels 102; Gaps 17;

QY 15 SEQVLRPALGSSVALNCT--AMVYSGPHCSLPYQWLKDGPLIGIGHYSLSHSWKYA 72
DB 559 SKQELQNPRLSTLWNLQNAFCVCEHQSF--LQWVDDORLLTVGA----- 604
QY 73 NSEVLVSVLGAVNTSTEVYGAFTCSIONISFSSFTLRAGPTSHVAVLAVLAL 132
DB 605 --EQMCAEPLDME-----DMVLSFRNATQLSKITIISVYV---TVLVS 646
QY 133 LLAALLVYKCRNLVLLYQDAYGEVEINDKLYDAVYSYSDCEPKRFV--NFLKPOLER 191
```

```

Db 647 VGVGLVVKFYFHLMLLCKCKYRGGE-----SIIDAFVIYSS--QDDDWNRNEIVKLEEG 700
Oy 192 RRGKFLDRLDRLPRAPESADLI-VNLSCKRLIYVLSADLSRWGSHSF-----242
Db 701 VPPOLCLHRDFTPGVAIANITIGCFHRSRKVIYVSOHFIOSRWCIFEXEIAQTWOF 760
Oy 243 ---REGICRLLELTRRPDIFTFEGQRPDPAPALRLRQHRMLV-----TLLMRPGSV 293
Db 761 LSSAGI-----IFVLCKLEKS-----LLRQVELYRLLSRNTYLEWE-DSV 802
Oy 294 TPSSDFKEVOLAPRKRYRYPVEGDPOTQLQDDKPMILRGRVPEGRLSEVDP 350
Db 803 LGRHVFWRRLRKAL-----LAGKPOS-----PEGTA-DAETNP 834

```

RESULT 6

```

ID TLR2_MOUSE STANDARD: PRT: 784 AA.
AC OQUNT: 09DEC4:
DP 01-MAR-2002 (Rel. 41, Created)
F 01-MAR-2002 (Rel. 41, Last sequence update)
L 01-MAR-2002 (Rel. 41, Last annotation update)
DE TOLL-like receptor 2 precursor.
GN TLR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND MUTAGENESIS OF PRO-681.
RX MEDLINE=20014145; Pubmed=10548109;
RA Underhill D.M., Ozinsky A., Hajjar A.M., Stevens A., Wilson C.B.,
RA Bassetti M., Aderem A.;
RT "The TOLL-like receptor 2 is recruited to macrophage phagosomes and
RL discriminates between pathogens.";
RL Nature 401:811-815(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Macrophage;
RX MEDLINE=99288048; Pubmed=10358136;
RA Heine H., Kirschning C.J., Lien E., Monks B.G., Rothe M.,
RA Golshock D.T.;
RT "Cutting edge: cells that carry a null allele for TOLL-like receptor 2
RL are capable of responding to endotoxin.";
RL J. Immunol. 162:6971-6975(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Macrophage;
RX Pubmed=10666214;
RA Matsuyuchi T., Takagi K., Musikacharen T., Yoshikai Y.;
RT "Gene expressions of lipopolysaccharide receptors, TOLL-like
RL 2 and 4, are differentially regulated in mouse T lymphocytes.";
RL Blood 95:1378-1385(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Adipocyte;
RX Pubmed=10823826;
RA Lin Y., Lee H., Berg A.H., Lisanti M.P., Shapiro L., Scherer P.E.;
RT "The lipopolysaccharide-activated TOLL-like receptor (TLR)-4 induces
RL synthesis of the closely related receptor TLR-2 in adipocytes.";
RL J. Biol. Chem. 273:24255-24263(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=21085660; Pubmed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Izawa M., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Balajov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Knehl P., Lewis S., Matsuo Y., Nakaido I., Pesole G., Quackenbush J.,

```

```

RA Schriani L.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barish G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kaniya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Wittaker C., Wilmink L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohlsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -1- FUNCTION: Cooperates with MD-2 and TIR6 to mediate the innate
CC immune response to bacterial lipoproteins and other microbial cell
CC wall components. Acts via MyD88 and TRAF6 leading to NF-kappa-B
CC activation, cytokine secretion and the inflammatory response. May
CC also promote apoptosis in response to lipoproteins (By
CC similarity).
CC -1- SUBUNIT: Binds MD-2 and TIR6 via the extracellular domain. Binds
CC MyD88 via their respective TIR domains (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Plasma membrane and
CC phagosome.
CC -1- TISSUE SPECIFICITY: Detected in a macrophage cell line, smooth
CC muscle, lung, spleen, thymus, brain and adipose tissue.
CC -1- PTM: N-glycosylated.
CC -1- SIMILARITY: BELONGS TO THE TOLL-LIKE RECEPTOR FAMILY.
CC -1- SIMILARITY: CONTAINS 1 TIR DOMAIN.
CC -1- SIMILARITY: CONTAINS 16 LEUCINE-RICH REPEATS (LRR).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL: AF185284; AAF04277.1; -
DR EMBL: AF124741; MAD6481.1; -
DR EMBL: AF216783; AAF28345.1; -
DR EMBL: AF165189; AAD48335.1; -
DR EMBL: AK005043; BAB23770.1; -
DR MGI: 1346060; Tlr2.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000483; LRR_Cterm.
DR InterPro: IPR003592; LRR_out.
DR InterPro: IPR000157; TIR.
DR Pfam: PF00560; LRR; 6.
DR Pfam: PF01463; LRRCT; 1.
DR PRINTS: PR00019; LEURICRPT.
DR SMART: SM00370; LRR; 1.
DR SMART: SM00082; LRRCT; 1.
DR SMART: SM00369; LRR_TYR; 2.
DR SMART: SM00255; TIR; 1.
DR PROSITE: PS0104; TIR; 1.
KW Receptor; Immune response; Inflammatory response; Signal;
KW Transmembrane; Repeat; Leucine-rich repeat; Glycoprotein.
FT SIGNAL 1 24
FT CHAIN 25 784
FT DOMAIN 25 587
FT TRANSMEM 588 608
FT DOMAIN 609 784
FT REPEAT 51 74
FT REPEAT 75 98
FT REPEAT 99 122
FT REPEAT 124 147
FT REPEAT 173 196
FT REPEAT 221 244
FT REPEAT 276 300
FT REPEAT 335 358
FT LRR 8.

```

FT REPEAT 359 384 LRR 9.
 FT REPEAT 386 411 LRR 10.
 FT REPEAT 412 436 LRR 11.
 FT REPEAT 438 456 LRR 12.
 FT REPEAT 457 476 LRR 13.
 FT REPEAT 477 500 LRR 14.
 FT REPEAT 502 521 LRR 15.
 FT REPEAT 522 548 LRR 16.
 FT DOMAIN 639 784 TIR.
 FT CARBOHYD 147 147 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 414 414 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT MUTAGEN 681 681 P->: ABOLISHES MYD88-BINDING AND RESPONSE TO MICROBIAL CELL WALL COMPONENTS.
 FT CONFLICT 59 59 I -> P (IN REF. 5).
 FT CONFLICT 82 82 I -> M (IN REF. 5).
 SO SEQUENCE 784 AA; 89448 MM; 606D56BF5F320A2 CAC64;
 TIR Match
 Best Local Similarity 6.7%; Score 144; DB 1; Length 784;
 Matches 74; Conservative 41; Mismatches 114; Indels 66; Gaps 12.
 QY 72 ANLSVLVSSVGLGVNVST-----EYVGA-----FTGSIONISFSST----- 109
 DB 495 ASLPEVLLVMKTRENAVSTPSKDLGSPFKLETTERAGNHFVCSCELLSFETETPALAOI 554
 QY 110 -----LORAGPT---SHVAVLA---LVLLALLLALLVYKCR 143
 DB 555 LVDPMDSVLCSPPRLHNRLODARPSVLECHQALVSGVCCALLLLILVGL-----CH 610
 QY 144 LVLLVW---ODAYGEVEINDGK-----LYDAYVSYSDCPEDRKRVNFILKQLERR-R 193
 DB 611 HFHGIMTVRMAMWAMLOAKRKPCKAPCRDVCYDAFYSE--ODSHWVENIMVQLENSDP 668
 QY 194 GYKLFLEDDRODLLPRAPREBADDLVNLSRCRRLLIVLSDAFISLAWMSHSPREDCULL-EL 252
 DB 669 PFKLCLEHRKRDVPGKWIIDNIIDISTEKSHKTVFVLESENFVSEWCKYELDFSHFRLFEDEN 728
 QY 253 TRRFIFTEGQRDRPAHPALRLRLQHSNHLVLLMPGVSVPSSDMEKVOAL 307
 DB 729 NDAALVLLLEPIERKAIPIQRCFKLRKIMNTKYLEW-PLDEQOQEVFVNLRTAI 782
 RESULT 7
 TLR4.PAPAN
 A TLR4.PAPAN STANDARD; PRT: 826 AA.
 I Q9TSP2:
 DB 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Toll-like receptor 4 precursor.
 GN TLR4.
 OS Papio anubis (Olive baboon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 CC Cercopithecoidea; Papio.
 NCBI_Taxid=9555;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=11104518;
 RA Smirnova I., Poltorak A., Chan E.K.L., McBride C., Beutler B.;
 RT "Phylogenetic variation and polymorphism at the Toll-like receptor 4 locus (TLR4).";
 RL (In) Genome Biol. 1:RESEARCH002.1-2.10(2000).
 CC -i- FUNCTION: Cooperates with MD-2 and CD14 to mediate the innate immune response to bacterial lipopolysaccharide (LPS). Acts via MYD88, TRAP and TRAF6, leading to NF-kappa-B activation, cytokine secretion and the inflammatory response (By similarity).
 CC -i- SUBUNIT: Belongs to the lipopolysaccharide (LPS) receptor, a multi-protein complex containing at least CD14, MD-2 and TLR4.
 CC Binds MD-2 via the extracellular domain. Binds MYD88 and TRAP via their respective TIR domains (By similarity).

CC	-1-	SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC	-1-	SIMILARITY: BELONGS TO THE TOLL-LIKE RECEPTOR FAMILY.
CC	-1-	SIMILARITY: CONTAINS 1 TIR DOMAIN.
CC	-1-	SIMILARITY: CONTAINS 20 LEUCINE-RICH REPEATS (LRR).
CC		
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.ch/announce/isb-sib.ch).
CC		or send an email to license@isb-sib.ch .
DR	EMBL; AF180964;	AARF07059.1; .
DR	EMBL; AF180962;	AARF07059.1; JOINED.
DR	EMBL; AF180963;	AARF07059.1; JOINED.
DR	InterPro; IPRO001611;	LRR.
DR	InterPro; IPRO000483;	LRR_cterm.
DR	InterPro; IPRO003592;	LRR_cut.
DR	InterPro; IPRO003591;	LRR_typ.
DR	InterPro; IPRO00157;	TIR.
DR	Pfam; PF00560;	LRR; 9.
DR	Pfam; PF01463;	LRRcm; 1.
DR	Pfam; PF01582;	TIR; 1.
DR	PRINTS; PRO00019;	LEUICHRPT.
DR	SMART; SM00370;	LRR; 1.
DR	SMART; SM00082;	LRRCT; 1.
DR	SMART; SM00369;	LRR_typ; 2.
DR	SMART; SM00255;	TIR; 1.
DR	PROSITE; PSS0104;	TIR; 1.
KW	Receptor; Immune response; Inflammatory response; Signal transmembrane; Repeat; Leucine-rich repeat; Glycoprotein.	
FT	CHAIN SIGNAL	1..23
FT	DOMAIN	24..631
FT	TRANSSEM	632..652
FT	DOMAIN	653..826
FT	REPEAT	53..76
FT	REPEAT	77..100
FT	REPEAT	101..124
FT	REPEAT	128..149
FT	REPEAT	150..173
FT	REPEAT	174..197
FT	REPEAT	203..225
FT	REPEAT	228..252
FT	REPEAT	277..303
FT	REPEAT	327..350
FT	REPEAT	351..372
FT	REPEAT	373..398
FT	REPEAT	400..421
FT	REPEAT	422..445
FT	REPEAT	447..469
FT	REPEAT	470..494
FT	REPEAT	495..518
FT	REPEAT	520..541
FT	REPEAT	543..569
FT	REPEAT	571..592
FT	DOMAIN	672..818
FT	CARBOHYD	35..35
FT	CARBOHYD	173..173
FT	CARBOHYD	205..205
FT	CARBOHYD	282..282
FT	CARBOHYD	309..309
FT	CARBOHYD	497..497
FT	CARBOHYD	526..526
FT	CARBOHYD	575..575
FT	CARBOHYD	624..624
FT	CARBOHYD	630..630
SO	SEQUENCE	826 AA; 94678 MW; 422777318E5F1769 CRC64;
Query Match	Best Local Similarity	6.6%; Score 142.5; DB 1; Length 826;
Matches	74; Conservative	22.8%; Pred. No. 0.00025;
		Mismatches 112; Indels 85; Gaps 14;

01-MAR-2002 (rel. 41, last annotation update)
 Toll-like receptor 4 precursor.
 TLR4.
 OS Pan paniscus (Pygmy chimpanzee) (Bonobo).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
 NCBI_TaxID=9597;
 [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=11104518;
 RA Smithova I., Poltorak A., Chan E.K.L., McBride C., Beutler B.;
 RT Phylogenetic variation and polymorphism at the Toll-like receptor 4
 locus (TLR4).
 RT (TLR4).
 RL (in) Genome Biol. 1:RESEARCH002.1-2.10(2000).
 CC -1- FUNCTION: Cooperates with MD-2 and CD14 to mediate the innate
 immune response to bacterial lipopolysaccharide (LPS). Acts via
 MyD88, TIRAP and TRAF6, leading to NF-kappa-B activation, cytokine
 secretion and the inflammatory response (By similarity).
 CC -1- SUBUNIT: Belongs to the lipopolysaccharide (LPS) receptor, a
 multi-protein complex containing at least CD14, MD-2 and TLR4.
 CC Binds MD-2 via the extracellular domain. Binds MyD88 and TIRAP via
 their respective TIR domains (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE TOLL-LIKE RECEPTOR FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 TIR DOMAIN.
 CC -1- SIMILARITY: CONTAINS 21 LEUCINE-RICH REPEATS (LRR).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 DR EMBL: AF179220; AAF05320.1; -.
 DR EMBL: AF179218; AAF05320.1; JOINED.
 DR EMBL: AF179219; AAF05320.1; JOINED.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000483; LRR_Cterm.
 DR InterPro: IPR003592; LRR_out.
 DR InterPro: IPR003591; LRR_Typ.
 DR Pfam: PFO00157; TIR.
 DR Pfam: PFO0560; LRR: 9.
 DR Pfam: PFO1463; LRRCT: 1.
 DR Pfam: PFO1582; TIR: 1.
 DR PRINTS: PRO00019; LEURICHRPT.
 DR SMART: SM00370; LRR: 2.
 DR SMART: SM00082; LRRCT: 1.
 DR SMART: SM00369; LRR_Typ: 2.
 DR SMART: SM00255; TIR: 1.
 DR PROSITE: PS50104; TIR: 1.
 DR Receptor: Immune response; Inflammatory response; Signal;
 KW Transmembrane; Repeat; Leucine-rich repeat; Glycoprotein.
 FT SIGNAL 1 23
 FT CHAIN 24 839
 FT DOMAIN 24 631
 FT TRANSMEM 632 652
 FT DOMAIN 653 839
 FT REPEAT 52 76
 FT REPEAT 77 100
 FT REPEAT 101 124
 FT REPEAT 128 149
 FT REPEAT 150 173
 FT REPEAT 174 197
 FT REPEAT 203 225
 FT REPEAT 228 252
 FT REPEAT 277 303
 FT REPEAT 307 330
 FT REPEAT 332 350
 FT REPEAT 351 372
 FT REPEAT 373 398
 FT REPEAT 400 421

FT REPEAT 422 445 LRR 15.
 FT REPEAT 447 469 LRR 16.
 FT REPEAT 470 494 LRR 17.
 FT REPEAT 495 518 LRR 18.
 FT REPEAT 520 541 LRR 19.
 FT REPEAT 543 566 LRR 20.
 FT REPEAT 568 592 LRR 21.
 FT DOMAIN 672 818
 FT CARBOHYD 35 35
 FT CARBOHYD 173 173
 FT CARBOHYD 205 205
 FT CARBOHYD 282 282
 FT CARBOHYD 309 309
 FT CARBOHYD 497 497
 FT CARBOHYD 526 526
 FT CARBOHYD 575 575
 FT CARBOHYD 624 624
 FT CARBOHYD 630 630
 SQ SEQUENCE 839 AA; 95637 MW; 3B328C5682127D37 CRC64;
 Query Match 5.2%; Score 134; DB 1; Length 839;
 Best Local Similarity 22.7%; Pred. No. 0.0014;
 Matches 78; Conservative 59; Mismatches 116; Indels 91; Gaps 15;
 11 FLSPSEDOVLRPALGSSVALNCTAWVSGPHCSLPYQWIKDGLPIGIGHSLEHSW 70
 516 FNLSSIQVLMHSHNFFSLD-----TPYKCLNSIQVL-----DYLNIHMTS 559
 QY 11 KANLEVLVSSVGVNTSTEVYGAFTCSIONISFSF-----TIQAGPTS- 117
 560 KQELQHPFSLAFLNLTQND-----FACCEHQSFQWIKDQQLVEVERMCCAPPSDK 615
 QY 118 -----HAAVASLVLALLLALLVYKCLANVLWQDYGVEVINDCK 163
 616 QGMPVLSLNTTCMNKTTIGSVLSVLSVAVLVYKFFHML-----LAGCIKGRGE 671
 QY 164 -LYDAVYSYSDCEDEKRFV-NFLTKPOLERRRGKFLDDRDLLPRAEPSADLL-VNLSR 220
 672 NIYDAVIYSS--QDEDWANELVKNLEGVPPFOLCHRDITPGVAIAANIHHGFFHK 729
 QY 221 GRRLIVLSDAFLSRWCSHSF-----REGICRLLELTRRPIFTFEGQRDPA 269
 730 SRKVIYVVSQHFQSRWCIFREYLAQTWQFLSSRAGI-----IFVLQKVER--- 776
 QY 270 HPALRLRLRHRHLV-----TLLMRGSGYTPSSDPFKEVQAL 307
 777 ----TLIRQVELYRLLSRNTYLEWE-DSVLGRHIFWRRLKAL 815
 RESULT 12
 ID TL22_CHICK STANDARD; PRT; 781 AA.
 AC Q9DGB6;
 DT 01-MAR-2002 (rel. 41, Created)
 DT 01-MAR-2002 (rel. 41, last sequence update)
 DE 01-MAR-2002 (rel. 41, last annotation update)
 GN TLR2-2.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX PubMed=11590137;
 RA Fukui A., Inoue N., Matsumoto M., Nomura M., Yamada K., Matsuda Y.,
 Toyoshima K., Seya T.;
 RT Molecular cloning and functional characterization of chicken
 Toll-like receptors, a single chicken Toll covers multiple molecular
 patterns.";
 RT J. Biol. Chem. 276:47143-47149(2001).

```

CC -! FUNCTION: Participates in the innate immune response to microbial
CC agents. Acts via MyD88 and TRAF6, leading to NF-kappa-B
CC activation, cytokine secretion and the inflammatory response.
CC Mediates the response to mycoplasma macrophage-activating
CC peptide-2kD (MALP-2).
CC -! SUBUNIT: Binds MyD88 via their respective TIR domains. Binds TR6
CC via their respective extracellular domains (by similarity).
CC -! SUBCELLULAR LOCATION: Type I membrane protein (by similarity).
CC -! TISSUE SPECIFICITY: Highly expressed in ovary. Also detected in
CC brain, heart, lung, liver, spleen and kidney, and at low levels in
CC placenta, muscle, testis and proventriculus.
CC -! PMW: N-glycosylated.
CC -! SIMILARITY: CONTAINS 1 TIR DOMAIN.
CC -! SIMILARITY: CONTAINS 13 LEUCINE-RICH REPEATS (LRR).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL: AB046533; BAB16842.1; -
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000483; LRR.
DR InterPro: IPR003592; LRR.
DR InterPro: IPR003591; LRR.
DR InterPro: IPR000157; TIR.
DR Pfam: PF00560; LRR.
DR Pfam: PF01463; LRRCT.
DR Pfam: PF01582; TIR.
DR PRINTS: PR00019; LEURICRPT.
DR SMART: SM00370; LRR.
DR SMART: SM00082; LRRCT.
DR SMART: SM00369; LRR.
DR SMART: SM00255; TIR.
DR PROSITE: PS50104; TIR.
DR Receptor: Immune response; Inflammatory response; Signal.
DR Transmembrane; Repeat; Leucine-rich repeat; Glycoprotein.
KW SIGNAL.
FT CHAIN 1 24
FT DOMAIN 25 781
FT TRANSMEM 585 585
FT DOMAIN 606 606
FT DOMAIN 607 781
FT REPEAT 51 74
FT REPEAT 75 98
FT REPEAT 100 122
FT REPEAT 124 147
FT REPEAT 148 171
FT REPEAT 173 195
FT REPEAT 195 356
FT REPEAT 356 381
FT REPEAT 383 409
FT REPEAT 410 433
FT REPEAT 435 453
FT REPEAT 454 473
FT REPEAT 474 497
FT REPEAT 499 518
FT DOMAIN 518 781
FT CARBOHD 37 109
FT CARBOHD 109 150
FT CARBOHD 150 184
FT CARBOHD 184 301
FT CARBOHD 301 313
FT CARBOHD 313 330
FT CARBOHD 330 390
FT CARBOHD 390 439
FT SEQUENCE 781 AA: 89094 MW: 721183996320454 CRC64;
Query Match Best Local Similarity 23.6%; Score 131; DB 1; Length 781;
Matches 60; Conservative 37; Mismatches 109; Indels 48; Gaps 9;

```

```

OY 96 FPCSTGNISFSSFTLORAG-----PTSHA-----AVLA 124
DB 532 FICSC---FISFHHNGEAGQVLGVWGESYICSPITRGAGVGYQLSLMECHRSLL 588
OY 125 SLVLALLALLAALLVYKRLNLLWYOD---AYGEVEINDK-----LYDAVYSDCP 175
DB 589 SLICHLVEFLFILLVVGKYKHAWYKMTWMLQAKRKPRAPTKDICYDAFVSYSE--646
OY 176 EKRKYVNTLKPQERR-RGKKFLDRDLPRAPESADLLVNSRCRLIVLSDAFLS 234
DB 647 NDSNWEVNIWQQLQEOACPRFLCLHKRDVPGKWIYNDISDEKSHKTLFVLSHFVQ 706
OY 235 RANCSHFRBEGICLL-ELTRRPITFEGORDPANRALRLRQNRHVTLLMPGGSV 293
DB 707 SHWKYEIDFHFRLDENNDVALILLLEPIQSQATPRKCKLRKIMNTKTYLEMPDE-765
OY 294 TPSSDFKREYQAL 307
DB 766 PQQMFENIKKAL 779
RESULT 13
TIR2_MACFA STANDARD; PRT; 784 AA.
AC OQ5M53;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE TIR2-like receptor 2 precursor.
GN TIR2.
OS Macaca fascicularis (Crib eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Monocytes;
RA Roberts F.A.; Tang S.;
RT "Macaca fascicularis Toll-like Receptor 2";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -! FUNCTION: Cooperates with MD-2 and TR6 to mediate the innate
CC immune response. Acts via MyD88 and TRAF6 leading to NF-kappa-B
CC activation, cytokine secretion and the inflammatory response. May
CC also promote apoptosis in response to lipopeptides (by
CC similarity).
CC -! SUBUNIT: Binds MD-2 and TR6 via the extracellular domain. Binds
CC MyD88 via their respective TIR domains (by similarity).
CC -! SUBCELLULAR LOCATION: Type I membrane protein (by similarity).
CC -! SIMILARITY: BELONGS TO THE TOLL-LIKE RECEPTOR FAMILY.
CC -! SIMILARITY: CONTAINS 1 TIR DOMAIN.
CC -! SIMILARITY: CONTAINS 14 LEUCINE-RICH REPEATS (LRR).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL: AY045573; AAK91868.1; -
DR PROSITE: PS50104; TIR.
DR Receptor: Immune response; Inflammatory response; Signal.
DR Transmembrane; Repeat; Leucine-rich repeat; Glycoprotein.
KW SIGNAL.
FT CHAIN 1 18
FT DOMAIN 19 784
FT TRANSMEM 588 587
FT DOMAIN 609 784
FT REPEAT 51 74
LRR 1.

```


QY 96 PFCIONISSFTLORAG-----PISHVA-----AVIA 124
 DB 544 FISCSE---FLSFTHHAGIAQVGVGPESYIDSPITVNGAOGVOSIMECHSLV 600
 QY 125 SLVLALLAALLVYKCRINVLWYOD---AYGEVEINDGK-----LYDAYVSYDCP 175
 DB 601 SLICTVLEFLILLVVGYKYHAWYMTWMLQAKRKFRAPTRDIDCDAFVSYSE-- 658
 QY 176 EDRKFVNFILKPOLERR-NGYKFLFDDRDLLPRAEPSADLVNLSRCRLIVLSDAFLS 234
 DB 659 NDSNMVENVIMVQOLEQACPFRLCLHNRDVPFGKWIYDNIIDISKSHKTLFVLSHFVQ 718
 QY 235 RANCSHSEFREGRL-ELFRRPFIETFEQGRDPANPALRLRHRHLVTLLMPGVS 293
 DB 719 SEMCKYELDSHRLFDENNDAVILLLEPIQSOAIPKRCRKLKIMNTKTYLEMPRDE- 777
 QY 294 TPSSDFWKEVOLAL 307
 DB 778 EDOOMENIKAL 791

TLR4_HORSE
 ID TLR4_HORSE STANDARD: PRT: 843 AA.
 AC G9WV3;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DE TOLL-like receptor 4 precursor.
 GN TLR4.
 OS Equus caballus (Horse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 NCBI_TaxID=9796;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Vandenaslas M.H., McNeill B.W., Barton M.H., Moore J.N.:
 RT Cloning and sequencing of equine toll-like receptor 4 (TLR4)."
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Cooperates with MD-2 and CD14 to mediate the innate
 immune response to bacterial lipopolysaccharide (LPS). Acts via
 MYD88, TIRAP and TRAF6, leading to NF-kappa-B activation, cytokine
 secretion and the inflammatory response. (By similarity).
 CC -1- SUBUNIT: Belongs to the lipopolysaccharide (LPS) receptor, a
 multi protein complex containing at least CD14, MD-2 and TLR4.
 CC Binds MD-2 via the extracellular domain. Binds MYD88 and TIRAP via
 their respective TIR domains (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE TOLL-LIKE RECEPTOR FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 TIR DOMAIN.
 CC -1- SIMILARITY: CONTAINS 19 LEUCINE-RICH REPEATS (LRR).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: AY005808; AAF91076.1; -
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000483; LRR_Cterm.
 DR InterPro: IPR003592; LRR_out.
 DR InterPro: IPR003591; LRR_Typ.
 DR InterPro: IPR000157; TIR.
 DR Pfam: PF00560; LRR_11.
 DR Pfam: PF01463; LRRCT_1.
 DR Pfam: PF01582; TIR_1.
 DR PRINTS: PR00019; LEURICHRPT.
 DR SMART: SM00370; LRR; 3.
 DR SMART: SM00082; LRRCT; 1.

DR SMART: SM00369; LRR_Typ; 1.
 DR SMART: SM00255; TIR; 1.
 DR PROSITE: PS50104; TIR; 1.
 KW Receptor; Immune response; Inflammatory response; Signal;
 KW Transmembrane; Repeat; Leucine-rich repeat; Glycoprotein.
 FT SIGNAL 1 23
 FT CHAIN 24 843
 FT DOMAIN 24 633
 FT TRANSMEM 634 654
 FT DOMAIN 655 843
 FT REPEAT 53 76
 FT REPEAT 77 100
 FT REPEAT 101 124
 FT REPEAT 128 149
 FT REPEAT 150 173
 FT REPEAT 174 197
 FT REPEAT 203 225
 FT REPEAT 277 300
 FT REPEAT 311 334
 FT REPEAT 351 373
 FT REPEAT 374 399
 FT REPEAT 401 422
 FT REPEAT 423 446
 FT REPEAT 447 470
 FT REPEAT 471 495
 FT REPEAT 496 519
 FT REPEAT 521 544
 FT REPEAT 545 566
 FT REPEAT 569 593
 FT DOMAIN 674 820
 FT CARBOHYD 35 35
 FT CARBOHYD 189 189
 FT CARBOHYD 205 205
 FT CARBOHYD 282 282
 FT CARBOHYD 295 295
 FT CARBOHYD 498 498
 FT CARBOHYD 527 527
 FT CARBOHYD 576 576
 FT CARBOHYD 626 626
 FT SEQUENCE 843 AA; 96495 MW; B597689AD677A69 CRC64;
 Query Match Best Local Similarity 22.08; Score 129; DB 1; Length 843;
 Matches 76; Conservative 49; Mismatches 113; Indels 108; Gaps 17;
 QY 42 CSLPSVQWIK-----DGLPLGIGGHYSLHEYSWKANLSEVLY-----SSV 82
 DB 518 CLPRLRLVLMNSHNNLLFLDMLP-----YKPLHSLLQIDCSFNRIVAFKQOELQHPSSL 572
 QY 83 LGVNVSTVEYVGAFTCSIONISF-----SFTLORAG-----PT 116
 DB 573 ASLINTQND---FACVCEYOSFLOVWQDORQLVEVEHLCALPLQMRGMPVLGFENNAT 628
 QY 117 SHVAAVLAS---LVLVLLAALLVYKCRINVLWYODAYGEVEINDGKLYDAYVSYSD 173
 DB 629 COISKTIYGGVSFSLMSVLAIVLYKFFFLMLAGCKKYGGE---SIYDAFVYSS 684
 QY 174 CEDERKFV-NFLIKQLERRRGYKFLFDDRDLLPRAEPSADLVNLSRCRLIVLSDA 231
 DB 665 -QDEDDWVRNELVKNLEBGPVPFQCLHRYDFITGVAAIANIIEGHHKSKRYIVVSSH 742
 QY 232 FLSTRMCSHSF-----REGICRLLELTRRPFIETFEQGRDPANPALRLRHR 280
 DB 743 FLQSRMCLFEYELIAQTWQFLSSRAGI-----IFIYVHKLEKS-----LLRQV 785
 QY 261 HLV-----TLLMRPGSVTPSSDFWKEVOLALPRKRYRYPGSDP 320
 DB 786 ELYRLNRYTLEW-DSVLDGRHITFWRRLNAL-----LDGKP 822

Search completed: November 7, 2002, 18:50:18
 Job time : 29 secs

C:Accession: A36187; S06928
R:Sim's, J.E.; Acres, R.B.; Grubin, C.E.; McMahon, C.J.; Wignall, J.M.; March, C.J.; Dowd
Proc. Natl. Acad. Sci. U.S.A. 86, 8946-8950, 1989
A:Title: Cloning the interleukin 1 receptor from human T cells.
A:Reference number: A36187; MUID:90046906
A:Accession: A36187
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-569 <SIM>
A:Cross-references: GB:M20658
R:Chua, A.O.; Gubler, U.
Nucleic Acids Res. 17, 10114, 1989
A:Title: Sequence of the cDNA for the human fibroblast type interleukin-1 receptor.
A:Reference number: S06928; MUID:90098789
A:Accession: S06928
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-569 <CHU>
A:Cross-references: EMBL:A16896; NID:g333800; PIDN:CA934773.1; PID:g333801
A:Genetics:
A:Gene: GDB:IL1R1; IL1RA; D2S1473; IL1R
A:Cross-references: GDB:125254; OMIM:147810
A:Position: 2q12-2q12
A:Superfamily: Interleukin-1 receptor type I
A:Keywords: cytokine receptor; transmembrane protein

Query Match	9.88;	Score 209.5;	DB 2;	Length 569;
Best Local Similarity	23.78;	Pred. No. 1.5e-09;		
Matches	85;	Conservative	62;	Mismatches 126;
			Indels	85;
			Gaps	16;

```

0Y 7 RAPDFLESPDDVYLRA-----LSSVALNCSTAMVSGPHCLSPFVOK-----L 50
Db 211 RVLEFITLEENKRPTRVAVISPAENMETMEVDGSOJOLICN---VTGLSDIAVMKNGSVI 267
0Y 51 KDGJPGIGGSHLYHEYSVKANISEVYSSVLGVANTSTREYV-GAFGCSIONI-SFSSF 108
Db 268 DEDDPVYLGEDYXSVEN---PANKRSPRLITVLNISIESERFKKHFFTOFANKHTGIDAA 322
0Y 109 TLRAGGTS---HYAAVYASILVYLLALLAALLAYVCRLNYLWOD---YGVSEVND 166
Db 324 YLQILIPVYNFOKMHMIGICVLYLVI--YCSVFYIKFKFIDYILWYRDSOYFELPKASD 388
0Y 162 GKXYDAVYVY-----SQCPEDKRFVNFILKPOLERRRGYKFLFDRDLRLPAEPESA 212
Db 382 GKXYDAVYILPKYVGEBSGSDOD---IFVEKYLPVEYLKOCGKILFYGRDYGVGEDI 438
0Y 213 DLYVNLSSRCRLIYVLSDAFLSRAMCHSFREGICR-----LLEL-----TR 254
Db 439 VYENENKRSKRLIILVRETSGLFSGLGSSBOJAMYNALYODGIKVYLLELEKIDOEYK 496
0Y 255 RPIJTFEGORQDPAPHALRLRLROHRLVYLLM---RFGSVYTPSSDFWEKVOALP 308
Db 499 MPSPKIFFIKOK---HGAIK-----WSDGFTGQDPSAKTRFKMKNVYRNM 539

```

```

RESULT 3
G02426
Interleukin-1 receptor-related protein - human
C/Species: Homo sapiens (man)
C/Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 29-Sep-1999
C/Accession: G02426
R/loverberg, T.W.; Crowe, P.D.; Liu, C.; Chalmers, D.T.; Liu, X.
submitted to the EMBL Data Library, February 1996
A/Reference number: H01239
A/Accession: G02426
A/status: preliminary; translated from GB/EMBL/DBD
A/molecule type: mRNA
A/residues: 1562<UD>
A/cross-references: EMBL:U49065; NID:g1236078; PIDN:AB553237.1; PID:g123607
A/superfamily: Interleukin-1 receptor type I

```

Query Match	9.58;	Score 204;	DB 2;	Length 562;
Best Local Similarity	28.58;	Pred. No. 4.3e-09;		

```

Matches      81:  Conservative      45:  Mismatches      110:  Indels      48:  Gaps      15:

QY      57  GIGGHSYLHXYSWKANKLSEVLVSSVLGVNTSYEV-----YG-AFCSIONISFSSFTIQ 111
Db      283  GVETHSFRKH-----:::|||||:::||||| 328
QY      112  RAGFTSHVAAVLASLLVLLALLAAL-LYKVCRLNVLMTQDAYGEVE-INDKLYDAYV 169
Db      329  LPADDER-ALLICGLALVAVAVSYYVYIYNFKIDYILWRSFAHSTETIVDKLTDAYV 387
QY      170  SYSDFPDRK-----FVNFILKPOLERRRGKYLEDDRDILPAEFSADLVNLSR 220
Db      388  LY---PRPHESGRHVAVDALVNLILPEYLHOCGKFLFISGRDFFPQQAANYIDENVXL 444
QY      221  CRRLLIVLSNAFLSRMCWCHSFREGLC---RLLELTRPRPIFTFEGORRPNAP-ALRL 276
Db      445  CRRLLIVLPVESIGFCLLNKLNSEQIAYVSAIDOGMKVILLIELEKIEDTYVMPESIOYI 504
QY      277  RQHRHLVTLLLMRPGSVTPSSD-----FMKEVOLALPRKYRPR 315
Db      505  KOKHGAIR---WH-GDFTQSOCKMTKFKMKTIVRYHMPPR-RCRP 543

RESULT 4
A32604
interleukin-1 receptor precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 29-Sep-1999
C:Accession: A32604
R:SJMS, J.E.; March, C.J.; Cosman, D.; Widmar, M.B.; Macdonald, H.R.; McMahon, C.J.;
S.K.

```

A: Amino acid: **Asn**
 A: Amino acid position: **14**
 A: Amino acid type: **Asn**
 A: Accession: **A32604**
 A: Molecule type: **mRNA**
 A: Residues: **1-576** <SIM>
 A: Cross-references: **GB:M20658; GB:M29752; NID:9198300; PIDD:AAA39279.1; PID:9309399**
 C: Superfamily: **Interleukin-1 receptor type I**
 C: Keywords: **Cytokine receptor; glycoprotein; phosphoprotein; transmembrane protein**
 F:1-19/Domain: **signal sequence** #status predicted <SIG>
 F:20-576/Product: **Interleukin-1 receptor** #status predicted <INT>
 F:339-359/Domain: **transmembrane** #status predicted <TM>
 F:46-99,145-199,251-315/Disulfide bonds: **#status predicted**
 F:63,103,114,236,257,266,300/Binding site: **carbohydrate (Asn) (covalently #status predicted**

Query Match	8.9%	Score 190.5;	DB 2;	Length 576;
Best Local Similarity	24.0%	Pred. No. 6e-08;		
Matches	86;	Conservative	71;	Mismatches 151;
				Indels 51;
				Gaps 18;

Oy	6	DRADPFLSPEDDOLVLPALGSSVALNCTAMVAVSSPHGLSPYOMKICPLGICGYSI#H	65
Db	227	DR-FVILSP-RNETIBADGSMIOLCH--VTGQFSDL--VYWKNGSGEIMNDPEFLAE	27
Oy	66	EYSWVK--AALSLVLSVSLGVAVN-STEVYG-AFTGSLONIS-FSSPTLORAGPSSHVA	12
Db	280	DYQVEHPSTKRRKTYLTITLNTSEVKSQFYRPLCYVKNNTNFEASHQOLIPYVDFEN	33
Oy	122	VIASLVLV--ALLAALLYVCGRIWLVNMOD--XGEVINDGKIVAY-----	16
Db	340	YLIGGFIILATIVCCVCATYKFKVDIYLMKTRDSCSGFLSKASDGTIDATILYPKLTG	39
Oy	170	--SYSDCPEDRKYVNTILKPOLERRRGCKFLPDDRLLPRAEFSADLLVYLSKRCRLIYV	22
Db	400	EGSSTDLDD--TFVKELDPVLEGGPGLKTLVYGDVDYVGDITLEVYENWVKSRRLIIT	45
Oy	228	LSDAFLSRAWCHSHSFREGICRLLELTRRPIFYT--EGQRDPAPHALRLILROHRLY	28
Db	457	LVYRDMGFSWLGOSSEBOAIYANALIGICITVLELEKIDYDEKIMDSITQIKQ-KHEV	51
Oy	284	TLLIIV-----RPGSVYTPSSDFWKEVOLAPRKVR-----YRFEVGDPOIQLDDDKP#	331
Db	516	-TQWSDQFOERQSA--KTRFKMKILRQMAQGRSFLSKHRIITLITDP--VADTRKKI	567

RESULT 5

Lymphocyte antigen Ly84 precursor - mouse

N:Alternate names: 38.5K T1 glycoprotein; ST2L protein

C:Species: Mus musculus (house mouse)

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 01-Dec-2000

C:Accession: S29498; A33541; S17657; S07054

R:Yamagisawa, K.; Takagi, T.; Tsukamoto, T.; Tetsuka, T.; Tomioka, S.
FEBS Lett. 318, 83-87, 1993A:Title: Presence of a novel primary response gene ST2L, encoding a product highly similar
A:Reference number: S29498; MUID:93170492

A:Accession: S29498

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-567 <FAN>

A:Cross-references: EMBL:D13695; NID:g286100; PIDN:BAA02854.1; PID:g286101

R:Lenz, R.; Hoffmann, S.; Wernskold, A.K.
Proc. Natl. Acad. Sci. U.S.A. 86, 5708-5712, 1989

A:Title: Serum- and oncoprotein-mediated induction of a gene with sequence similarity to

A:Reference number: A33541; MUID:89345536

A:Accession: A33541

A:Molecule type: mRNA

A:Residues: 1-191, 'V', 193-328, 'SKECPSHIA' <KLE>

A:Cross-references: GB:M24843; NID:g201103; PIDN:AAA0160.1; PID:g201104

R:Tominaga, S.I.; Jenkins, N.A.; Gilbert, D.J.; Copeland, N.G.; Tetsuka, T.
Biochim. Biophys. Acta 1090, 1-8, 1991

A:Title: Molecular cloning of the murine ST2 gene. Characterization and chromosomal map

A:Reference number: S17657; MUID:91355215

A:Accession: S17657

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-191, 'V', 193-328, 'SKECPSHIA' <TKM>

A:Cross-references: EMBL:X60184; NID:g54200; PIDN:CAA42742.1; PID:g54201

R:Tominaga, S.I.
FEBS Lett. 258, 301-304, 1989

A:Title: A putative protein of a growth specific cDNA from BALB/c-3T3 cells is highly si

A:Reference number: S07054; MUID:90092495

A:Accession: S07054

A:Molecule type: mRNA

A:Residues: 1-328, 'SKECPSHIA' <TQ2>

A:Cross-references: EMBL:X07519; NID:g55517; PIDN:CAA68812.1; PID:g55518

A:Note: It is uncertain whether Met-1, Met-7 or Met-19 is the initiator

C:Genetics:

A:Gene: ST2

A:Position: 1

A:Exons: 27/1; 97/2; 155/3; 210/1; 233/1; 280/2

C:Superfamily: Interleukin-1 receptor type I

C:Keywords: glycoprotein

F:1-26/Domain: signal sequence #status predicted <STG>

F:27-337/Product: ST2 protein #status predicted <MAT>

F:60,101,107,146,176,194,225,259,278/Binding site: carbohydrate (Asn) (covalent) #status

Query Match Best Local Similarity 8.8%; Score 189.5; DB 2; Length 567;

Matches 82; Conservative 59; Mismatches 153; Indels 37; Gaps 14;

QY 9 PDLSPSEDQVLRPALGSSVALNCTAMVYSGPHSLPSYQWIKDKGLIGIGHYSLHEXS 68

Db 217 PVATNPVNHMTMEVEIGKPAISACFGKSH-FLADVLMQINKTVGNFGEARIQEE 275

QY 69 WVKANISEVL-VSSYLVG-VNTSTEV---YGAFTCSIONISSSTLORAGTSH-VA 120

Db 276 GRNESSNDMDCTSVLRGTGTEKDSLSEYDCLALNLHGMIKRTLRKQPIIDHRSTY 335

QY 121 AVLASLVLLALLALLVK-C-RLNVLWYODAYGEVEI-NDGKLYDAYVSY----- 171

Db 336 YIVAGSLLMLFIVLVYLVKFWIEVALFWRDIVPYTRNDGKLYDAYIIPRFGRS 395

QY 172 -SDCPEDRKVFVILKPOLERRRGYKLLDRLDLPRAEPSADLVNLSRCRLIYVLS 230

Db 396 AAGTHSVEYFVHHTLPDVLNKGKGYKLCITYGRDLLPGQAAAFVSSIONSRQVFLAP 455

QY 231 AFLSRWCSHSFREGI-CRLLETRRPITFTPEGQRDPANPALRL-----LRQH-RHLVT 284

Db 456 HMMHKEFAVEQDIALHSLIIONNSKVILLIEME-----PLGEASRLQGVDDQSLQHLVK 510

QY 285 L--LLMRGVSYP-----SSDFWEVOLALP 308

Db 511 IQGTIKWRDHDVADKQSLSKFWKHYRYQMP 541

RESULT 6

JQ1526

Interleukin-1 receptor I precursor - chicken

C:Species: Gallus gallus (chicken)

C:Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 29-Sep-1999

C:Accession: JQ1526

R:Guida, S.; Heguy, A.; Mell, M.
Gene 111, 239-243, 1992

A:Title: The chicken IL-1 receptor: differential evolution of the cytoplasmic and ext

A:Reference number: JQ1526; MUID:92175529

A:Accession: JQ1526

A:Molecule type: mRNA

A:Residues: 1-555 <GVU>

A:Cross-references: GB:M81846; NID:g212206; PIDN:AAA48924.1; PID:g212207

A:Experimental source: embryo

C:Superfamily: Interleukin-1 receptor type I

C:Keywords: cytokine receptor; transmembrane protein

F:1-13/Domain: signal sequence #status predicted <STG>

F:20-555/Product: Interleukin-1 receptor I #status predicted <INT>

F:339-359/Domain: transmembrane #status predicted <FRA>

Query Match Best Local Similarity 8.6%; Score 184; DB 2; Length 555;

Matches 90; Conservative 65; Mismatches 101; Indels 104; Gaps 21;

QY 9 PDLSPSEDQVLRPALGSSVALNCTAMVYSGPHSLPSYQWIKDKGLIGIGHYSLHEXS 52

Db 230 PERYTNP-NNTIEVLELGSYHVMCN--VSSGYGLLPYQVNDQEDVDSPSTYRBEFYEE 286

QY 53 GLPLIGIGHYSLHEYSWKANISEVLSSVLYGVNTSTEVYGAFTCSI-ONISFSSTLQ 111

Db 287 GMPHGIA-----VSGTFENISEVVKLDY-----AYKFFCHFIYDSQEFSTY-IR 329

QY 112 RAGPTSHVAVL-----ASLLVLLALLALLVKCKLANLWYQDAYGEV---EINDGKL 164

Db 330 LEHPVQINIGYLLGGISLIFLFLIL--IVYKIFIDIVLWYRSSCHPLKRVSDGKI 387

QY 165 YDAYVSYSDCPEDRK-----FVNFILKPOLERRRGYKFLDRLDLPRAEPSADL- 215

Db 388 YDAYVYX---PKNRSECLYSSDIFALKILPEVLERCQGYNLFIFGRNDL-AGEAVYDVT 443

QY 216 VNLSCRRLIVL-----SDAFLSRWCSHSFREGI-CRLLETR-----R 255

Db 444 EKIHQSRVYIILVPEPCYGLIEDASEKHL--AVYNALIQDGIKILLIELEKIDYANM 501

QY 256 PIPTFEGQRDPANHALUL-----RQHRHLVTLMLMRPGSTPSDWEKVEQLAPRK 310

Db 502 PESIKYVKOK---YGAIRMTGDFSERH-----SASTRWKRVRYRMP 543

RESULT 7

A29943

Toll protein precursor - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C:Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 02-Jun-2000

C:Accession: A29943

R:Hashimoto, C.; Hudson, K.L.; Anderson, K.V.
Cell 52, 269-279, 1988

A:Title: The Toll gene of Drosophila, required for dorsal-ventral embryonic polarity,

A:Reference number: A29943; MUID:88135760

A:Accession: A29943

A:Molecule type: DNA

A:Residues: 1-1097 <HMS>

A:Cross-references: GB:M19969; GB:J02682; NID:g158640; PIDN:AAA26941.1; PID:g158641
 C:Genetics:
 A:Gene: FlyBase:FLY
 A:Cross-references: FlyBase:FBgn0003717
 C:Superfamily: leucine-rich alpha-2-glycoprotein repeat homology
 F:1-17/Domain: signal sequence #status predicted <SIG>
 F:18-1097/Product: toll protein #status predicted <MAT>

Query Match
 Best Local Similarity 8.2%; Score 176; DB 2; Length 1097;
 Matches 70; Conservative 47; Mismatches 89; Indels 66; Gaps 13;

OY 121 AVATLVLALLALLL-----VKGR-----NVLWYQDAVEINDGLYDAYVSY 172
 DB 810 AVATLVLALLLALLL-----VKGR-----NVLWYQDAVEINDGLYDAYVSY 172
 OY 173 DCEDEKRFVNFILKPOLER-RGCVLFLLDDNDLLPRAEADLLVNLSCRLIVLSDA 866
 DB 867 H--KDQSFIEDLVLPQLEHGPOKQFQDCHVEDMDLVGHIPIENIRVADSRRITVLSQ 924
 OY 232 FLIRACMSHSP-----EGLCLLELTLRPIFTFEQGRDPAHALRLRHRHIV 283
 DB 925 FLIRACMSHSP-----EGLCLLELTLRPIFTFEQGRDPAHALRLRHRHIV 283
 OY 284 TLLWRPQSVPPSSD--FKREVQALAPRKVRYPEGDPOTQDQDPMILRGVPEGS 341
 DB 976 TYLKM-----GDPWMDKRLRFLHR--RV-----GNING 1005
 OY 342 RALDSEVDPECDLGVGKGVGPPSPAPHTS 373
 DB 1006 ALIKTALKGSTDDKLELKP--SPVPLPLT 1034

RESULT 8
 156526
 Interleukin 1 receptor type I - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 29-Sep-1999
 C:Accession: 156526
 J:Heart, R.P.; Liu, C.; Shadick, A.M.; McCormack, R.J.; Jonakait, G.M.
 A:Title: An mRNA homologous to interleukin-1 receptor type I is expressed in cultured rat
 A:Reference number: 156526; MUID:93266794
 A:Accession: 156526
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-390 <RES>
 A:Cross-references: GB:M95578; NID:g451305; PIDN:AAA16196.1; PID:g451306
 C:Superfamily: interleukin-1 receptor type I
 C:Words: cytokine receptor

Query Match
 Best Local Similarity 7.8%; Score 168.5; DB 2; Length 590;
 Matches 88; Conservative 56; Mismatches 128; Indels 83; Gaps 21;
 OY 6 DPAPFLSPSEDOVLRPALGSSVLANCTANVWSPHOSLPYQWLDGLPLGIGHYSUH 65
 DB 241 DR-PVINSR-PNEMTEADPGSTIOLCN--VIGQFIDL--VYMKNGSEIEMDDPIAE 293
 OY 66 EYSWVK--ANLSEIVSVYGVNTSTEVYQ-AFICSDIONISFSFTLQKRGPTSHVAA 121
 DB 294 DQFHEHSAARKYTLITL--LNWSEVSKOYRYPFICFVKNHILE-----TAHVRL 344
 OY 122 VLA-----SLVLLALLLALLLYKC-----RLNVLYQDAYGEV--EINDGLYDA 167
 DB 345 VYVPDPRKNVILGFAIFTAFCACIYKVRVDIVMYRSCSDPLPKASDGRYDA 404
 OY 168 YVSYSDCPEDEKRV--NFLTK--POLERRGYKFLDDRDLLRAPASDVLVNSRC 221
 DB 405 YVLPKPYGSGSPAYLDTFVFKLLPEVLEGQGYKFLCGRNDYGEPTIEVTENKRS 464
 OY 222 RLIVLVLSDAFLSRACSHS-----FRGL-CRIEL-----TRPPIFTFEG 263

DB 465 RLIIILVDMKSFSCLOSSFEQIAITDALIRGIIILILEKTIQDYKMPESIOFIK 524
 OY 264 QRRDPAPRLRLRQHRHLVLLLM-----RPSVPSDDFWKEVOLAPRKVR 312
 DB 525 QK-----HGA-----ICMSGDFKRFQSA--KTRFWKNRLRYQMPAQR 560

RESULT 9
 542633
 Fit-1M protein - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 25-Dec-1994 #sequence_revision 10-Nov-1995 #text_change 11-Jan-2000
 C:Accession: 542633
 R:Bergers, G.; Reikertorfer, A.; Braselmann, S.; Graninger, P.; Busslinger, M.
 A:Title: Alternative promoter usage of the Fos-responsive gene Fit-1 generates mRNA 1
 A:Reference number: 542632; MUID:94178260
 A:Accession: 542633
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-247 <BER>
 C:Superfamily: Interleukin-1 receptor type I

Query Match
 Best Local Similarity 7.6%; Score 163; DB 2; Length 247;
 Matches 65; Conservative 39; Mismatches 95; Indels 32; Gaps 11;

OY 110 LORAGPTSHVA--VLASLVLALLLALLLYKC--RLNVLYQDAYGEV--NDGKLY 165
 DB 3 LRRKQPIHQOSTYVYVAGCSLLMFINVLYLVKVFIEVALFMDIMADYKQNDGKLY 62
 OY 166 DAVSY-----SDCEDEKRFVNFILKPOLERRRGCVLFLLDDNDLLPRAEADLLV 217
 DB 63 DAVIYTPRVFNGSAGCTGSEVFVHYTLPQVLENGKCVKIQIIGRDLQDQDAYVSS 122
 OY 218 LSRCLRLIVLSDAPLIRACMSHSPFREG-CRIELRLRPIFTFEQGRDPAHALRL- 275
 DB 123 IONSROYFVLAIPHMHKKEFAVEQETALSHALIONNSKYLLEME-----FMGESRLQ 177
 OY 276 ---LRQH-RHLVTL--LWRFPGSVTP--SSDFKEVOLAPRKVRYP 315
 DB 178 LGDLDLSLQHLVKNQGTIKRQEDHVDKQSLSKFVKHVRQMP--VPRRP 226

RESULT 10
 713887
 tlr protein - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Nov-2000
 C:Accession: 713887
 R:Chiang, C.; Beachy, P.A.
 A:Title: Expression of a novel Toll-1-like gene spans the parasegment boundary and controls
 A:Reference number: 713887; MUID:95151581
 A:Accession: 713887
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1385 <CHI>
 A:Cross-references: EMBL:S76155; NID:g913247; PID:g913248; PIDN:AA033383.1
 C:Genetics:
 A:Cross-references: FlyBase:FBgn0004364

Query Match
 Best Local Similarity 6.7%; Score 144.5; DB 2; Length 1385;
 Matches 68; Conservative 44; Mismatches 104; Indels 45; Gaps 12;
 OY 99 STONTSFSSFLORAGPTSHVA--VLASLVLALLLALLLYKC--RLNVLYQDAYG-- 155
 DB 979 SASNTSSQ--DLAG--GRLPLAVALVILFDVVLIVFVRESVMMWLFAYGVRA 1033
 OY 156 EVELND-GKLYDAVSYSDCPEDEKRFVNFILKPOLER-RGCVKFLDDRDLLPRAEPA 212

```

Db 1034 CEPREFDAGKLYDAIILHSE--KDYEFCVCRNIAEILEHGHPRLICIQORDLPPQAS-HL 1090
QY 213 DILVINSRCRLIVLSDAFLSRAMCSHSFREGLCRLLE-LTRRPPIFI---TTEGQRDP 268
Db 1091 QLVGARASRKIIIVLTRNLATEMNRIEFRNAFHESLRGLAOKLVIIEETSVSAEADY 1150
QY 269 AH-----PALRLRQHRHLVTLMLMRGVSYPSSDFKWEVOLAPRKV--RYRVEG 318
Db 1151 AELSPLKSPSRNRLTCDRY-----TWKLRVAPIELSPRGNNYTL 1193
QY 319 DPQOTQDDKDPMLILGRVP 339
Db 1194 DHHERFKOPVSPGMIFRQAPR 1214

RESULT 11
11 2
g:Species: Drosophila melanogaster
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Nov-2000
C:Accession: T13852
R:Elidon, E.; Kooyer, S.; D'aveelyn, D.; Duman, M.; Lawinger, P.; Botas, J.; Belien, H.
Development 120, 885-899, 1994
A:Title: The Drosophila 18 wheeler is required for morphogenesis and has striking simila
A:Reference number: 217796; PMID:95324375
A:Accession: T13852
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1389 <EID>
A:Cross-references: EMBL:L23171; NID:g415682; PID:g1019104; PIDN:AAAT9208.1
C:Genetics:
A:Gene: wheeler
A:Cross-references: FlyBase:FBgn0004364

Query Match 6.5%; Score 139.5; DB 2; Length 1389;
Best Local Similarity 23.2%; Pred. No. 0.0036;
Matches 66; Conservative 51; Mismatches 105; Indels 63; Gaps 12;

QY 75 SELVSSVGLVANTSTEVYGAFCSTIONISFSSFTIQRAGPSTHVAVALSLVLLALL 134
Db 973 SDLLDASASNIS--SSODLACAIIC-----PCWPAVLVLIPLV 1009
QY 135 AALLVYKRLNVLWYQDAG---EVEIND-GKLYDAYSVSDCPEDRKVFETLKPOL 189
Db 1010 VLIIVFRRSVMMLEPAHKGVAVCEPRDAGKLYDAIILHSE--KDYEFCVCRNIAEL 1067
QY 190 ER-RRGYKFLDRDRLIPRAEPSADLLVNLSCRLIVLSDAFLSRAMCSHSFREGLCR 248
Db 1068 EHGRPFRLICIQORDLPPQAS-HLQLVGARASRKIIIVLTRNLATEMNRIEFRNAFHE 1126
QY 249 LLE-LTRRPPIFI---TTEGQRDPAH-----PALRLRQHRHLVTLMLMRGVSYP 296
Db 1127 SLRGLAOKLVIIEETSVSAEADYAEVLSPLKSPSRNRLTCDRY-----TWKLRVAPIELSPRGNNYTL 1193
QY 297 SDFWKEVOLAPRKV--RYRVEGDPOTQDDKDPMLILGRVP 339
Db 1172 --TWKLRVAPIELSPRGNNYTLDHHERFKOPVSPGMIFRQAPR 1214

RESULT 12
type I interleukin-1 receptor - rat (fragment)
C:Species: Rattus sp. (rat)
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 13-Sep-1998
C:Accession: I51903
R:Sutherland, D.B.; Varilek, G.W.; Neill, G.A.
Am. J. Physiol. 266, C1198-C1203, 1994
A:Title: Identification and characterization of the rat intestinal epithelial cell (IEC-
A:Reference number: I51903; PMID:94262728
A:Accession: I51903
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA

```

```

A:Residues: 1-169 <RES>
A:Cross-references: GB:S70238; NID:g546976
A:Experimental source: intestinal epithelial IEC-18 cells
C:Superfamily: interleukin-1 receptor type I
C:Keywords: cytokine receptor

Query Match 6.0%; Score 128.5; DB 2; Length 169;
Best Local Similarity 28.1%; Pred. No. 0.002;
Matches 50; Conservative 29; Mismatches 76; Indels 23; Gaps 6;

QY 145 NWLLMYQDAYGEV---EINDGKLYDAYVSYSDCPEDRKRV---NFIK--PQ-LERRRY 195
Db 1 DIYLMYRDSCDFLPPKASDGYDAYVLYPKYVGESEFAYLDTEFVKLLPEVLEGGQFG 60
QY 196 KLFELDRDLIPRAEPSADLLVNLSCRLIVLSDAFLSRAMCSHSFREGLCRLLETRR 255
Db 61 KLFICGRDYYGEDTLEVTNENVKRSRRLIILVRDMGSEFCGSGSEIOAIYDALIRE 120
QY 256 PIRTF---EGQRDPAPRALRLRQHRHLVTLMLMRGVSYPSSDFKWEVOLAPR 309
Db 121 GIVTILLELEIODYKMEPSIOFIKQ-----KHGATCGSDFERKPOSATKR 168

RESULT 13
S18209
fibroblast growth factor receptor 4 precursor (clone 61) - mouse
N:Alternate names: tyrosine kinase Mpk-11
N:Contains: protein-tyrosine kinase (EC 2.7.1.112)
C:Species: Mus musculus (house mouse)
C:Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 16-Jul-1999
C:Accession: S18209; S26751; S30497
R:Stark, K.L.; McMahon, J.A.; McMahon, A.P.
Development 113, 641-651, 1991
A:Title: FGFR-4, a new member of the fibroblast growth factor receptor family, expres
A:Reference number: S18209; PMID:92146274
A:Accession: S18209
A:Molecule type: mRNA
A:Residues: 1-799 <STAI>
A:Cross-references: EMBL:X59927
R:Stark, K.L.
submitted to the EMBL Data Library, May 1991
A:Reference number: S26751
A:Accession: S26751
A:Molecule type: mRNA
A:Residues: 1-485, 'QVYRAAPG', 486-799 <STAZ>
A:Cross-references: EMBL:X59927; NID:g50966; PIDN:CAA42551.1; PID:g50966
R:Gibaldi-Hebensstreit, P.; Nieto, M.A.; Fraint, M.; Mattei, M.G.; Chestier, A.; Wilkin
Oncogene 7, 2499-2506, 1992
A:Title: An Eph-related receptor protein tyrosine kinase gene segmentally expressed i
A:Reference number: S30496; PMID:93096484
A:Accession: S30497
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 611-667 <GIL>
A:Cross-references: EMBL:X57236; NID:g53187; PIDN:CAA40512.1; PID:g53188
C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; pr
C:Keywords: ATP; duplication; glycoprotein; growth factor receptor; phosphotransferas
F:19-799/Product: fibroblast growth factor receptor 4 #status predicted <STO>
F:19-799/Product: extracellular #status predicted <EXT>
F:162-223/Domain: immunoglobulin homology <IMM>
F:367-387/Domain: transmembrane #status predicted <TM>
F:388-799/Domain: intracellular #status predicted <INT>
F:462-747/Domain: protein kinase homology <KIN>
F:470-478/Region: protein kinase ATP-binding motif
F:54-68,169-221,268-330/Disulfide bonds: #status predicted
F:500,517,609/Active site: Lys, Glu, Asp #status predicted

Query Match 5.7%; Score 123; DB 2; Length 799;
Best Local Similarity 22.1%; Pred. No. 0.043;
Matches 103; Conservative 45; Mismatches 136; Indels 182; Gaps 23;

QY 24 LGSSVALNCTAWVSPGHCISLPSVQWLK---DGLPLGIGGHSLSHEYSWVANKANSEVLV 79

```

```

Db 260 VGSDFVLLCVCVSDADPH-----IOMLKHYVINGSSFGADGFPYQVLTPTDINISEV-- 312
QY 80 SSVLYGNVSTVEYGAFTCSION---ISFSS-----FTLORAGP---TSHVAV 122
Db 313 -QVLYLNWVAEDAGEYTCAGNSIGLSQSAWLVLPEDLTMTATTEARATITIIIV 371
QY 123 LASLLVLLALLALIVKCRNLVLMYODAYGEVEINDKLYDAVVSDDCPEDRKFVN 182
Db 372 SGSLVLLVLLLVAGYHROY-----IRGHYSROPVTIOKL-----SRPLARQF-- 415
QY 183 FILTKPOLER-----RGYKFLDDRDLIPRAEPSADLLVNSRCRRLIVLSDAF 232
Db 416 -----SLESRSSGKSSLVLRGVRLSSGCPPLITG-----LVNL-----DLP 452
QY 233 LSRAMGSHFREGICRLLELTR-----RPI-----FITEGORRDPAPH----- 271
Db 453 LDDPW-----EFPDRVLVKGKPLGEGCGQVYRAEAFQMDSPRDQSTVAV 499
QY 272 -----ALRLRQHRHLVTLILMRPGSVTPSSDFWKEVOLALPRKY 311
Db 500 KMLKDNASDKDLADLVSEMEVWKILGRNKNITNL---GVCTQHGPLYIVTECAKGNL 555
QY 312 RVRPVGDPQOTQLODDKDPMLIRGRVPRGRALDSEVDPDPDGGLGVKQVYGEPSAPPH 371
Db 556 RE-----FLRARPPG-----PDLSPGDPSSSGPL---SPAL 586
QY 372 TS-----GVALGSRSESVSDLSGRNYSARKDFYCIYSKDM 410
Db 587 VSCAVQVARGMQYLESR--KCIHRDLAARN-----VLYTEDV 622

```

RESULT 14

```

S11226
Myd88 protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 18-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 05-Nov-1999
C:Accession: S11226; J02079
R:Lord, K.A.; Hoffman-Liebermann, B.; Liebermann, D.A.
Oncogene 5, 1095-1097, 1990
A:Title: Nucleotide sequence and expression of a cDNA encoding Myd88, a novel myeloid differentiation factor
A:Reference number: S11226; MUID:90326414
A:Accession: S11226
A:Molecule type: mRNA
A:Residues: 1-243 <LOR>
A:Cross-references: EMBL:X51397; NID:953293; PIDN:CAA35762.1; PID:953294
R:Hultmark, D.
Biochem. Biophys. Res. Commun. 199, 144-146, 1994
A:Title: Macrophage differentiation marker Myd88 is a member of the TOLL/TL-1 receptor family
A:Reference number: J02079; MUID:94168566
A:Notes: annotation

```

```

Query Match 5.4%; Score 115; DB 2; Length 243;
Best Local Similarity 29.2%; Pred. No. 0.044;
Matches 33; Conservative 26; Mismatches 38; Indels 16; Gaps 5;

QY 163 KLYAYVYSDDCPEDRKFVNFIKPOLERRRGYKFLDDRDLIPRA--EPSADLLVNS 219
Db 106 ELPAFFICY--CPNDIEFVQEMIROLEQTYRLKLCVSDRVLPGTCWSTIASLEIE-K 161

QY 220 RCRRLIVLSDAFLSRAMGSHFREGICRLLELT-----RRPIITFEGORRDP 267
Db 162 RCRRWVVVSDYIQSKCEDFOTKFA-----LSLSPGVQKRLPIKYYKAMKD 210

```

RESULT 15

```

T08664
Toll protein-like receptor DKFZp54710610.1 - human
C:Species: Homo sapiens (man)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 23-Jul-1999
C:Accession: T08664
R:Poustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
Submitted to the Protein Sequence Database, May 1999

```

```

A:Reference number: 216466
A:Accession: T08664
A:Molecule type: mRNA
A:Residues: 1-786 <POU>
A:Cross-references: EMBL:AL050262
A:Experimental source: fetal brain; clone DKFZp54710610
A:Note: DKFZp54710610.1

```

```

Query Match 5.3%; Score 114.5; DB 2; Length 786;
Best Local Similarity 22.8%; Pred. No. 0.22;
Matches 49; Conservative 39; Mismatches 88; Indels 39; Gaps 9;

```

```

QY 119 VAAVLASLVLLALLALIVKCRNLVLMYODAYGE-----VEINDGKLY 165
Db 583 IVTLVATMLVLAIVTSLICYLD-----LPWYLRWVCQMTQTRRRARNIPLEIQRNQG 637
QY 166 DAYVYSDDCPEDRKFVNFIKPOLERRRGYKFLDDRDLIPRAEPSADLLVNSRCRRLI 225
Db 638 HAFISYSG--HDSFWYKNELLPMLE--KEGMQICLHERNFVPGKSIENITICIEKSYKSI 694
QY 226 VVLSDAFLSRAMGSHS-----FREGICRLLELTRRPIITFEGORRDP-AHPALRL 276
Db 695 FVLSPNFVQSEWCHYELFYAHNMLFHGSGNSLTILILEPI-----PQYSIPSSYHKLSL 749
QY 277 ROHRHLVTLILMRPGSVTPSSDFWKEVOLALPRKY 311
Db 750 MARR--TYLEW--PKRSKRGFLWMLRAAINIKL 780

```

```

Search completed: November 7, 2002, 18:52:57
Job time : 49 secs

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2002, 16:47:01 ; Search time 1676 Seconds
(without alignments)
9929.442 Million cell updates/sec

Title: US-09-598-443-1
1333
Sequence: 1 atgcacaggtctctgtatag.....tgcacagatgatagtag 1233

Scoring table: IDENTIFY_NUC
Gapop 10.0 , Gapext 1.0

Sequenced: 13736207 seqs, 674847542 residues
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: em_estb1a:*
2: em_estbhum:*
3: em_estb1:*
4: em_estb2:*
5: em_estb3:*
6: em_estb4:*
7: em_estb5:*
8: em_estb6:*
9: gb_est1:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	733	59.4	763	10	BI517905 603041985
2	725.2	58.8	941	10	BG740350 602634105
3	681.4	55.3	821	10	BE797163 601587171
4	660.4	53.6	852	10	BI521062 603081652
5	639.2	51.8	749	10	BI821947 603039758
6	616	50.0	808	10	BI769076 603057053
7	606.6	49.2	791	10	BI518182 603041985
8	589.8	47.8	845	10	BE792803 601584677
9	556.8	45.2	941	10	BI518544 603061669
10	546.8	44.3	1067	10	BI752545 603021956
11	541	43.9	983	10	BI759065 603042860
12	535.8	43.5	796	10	BG742077 602633540
13	522.4	42.4	801	10	BI770099 603053464
14	513	41.6	783	10	BI820133 603037143
15	496.4	40.3	821	10	BI758455 603022612
16	488.4	39.6	513	10	BF239133 601905529
17	486	39.4	925	10	BI905958 603062841

18	482.2	39.1	970	10	BE250059 600943075
19	468.8	38.0	822	10	BI909088 603070070
20	467.8	37.9	892	10	BE612529 601452037
21	454.8	36.9	458	9	AV655410
22	452	36.7	463	10	BE304820 601143635
23	452	36.7	878	10	BI103498 601157924
24	446	36.2	627	10	BE279363 601157924
25	435	35.3	540	9	AW786417 119555 MA
26	434.4	35.2	773	10	BI837418 603086710
27	427.6	34.7	1021	10	BI519672 603061669
28	418.8	34.0	705	10	BE869865 601446507
29	411	33.3	565	9	BE293826 601186875
30	402	32.6	567	9	AW250249 2821373.5
31	399.2	32.4	536	10	BG609794 323769 MA
32	390	31.6	714	10	BI101849 602885412
33	377.2	30.6	410	9	AA310938 EST181714
34	366	29.7	822	10	BE869379 601445316
35	364.8	29.6	822	10	BI838735 603082175
36	357.4	29.0	720	10	BF971860 602240449
37	353.4	28.7	629	10	BG333328 602431365
38	350.6	28.4	863	10	BE293423 600943075
39	343.8	27.9	698	10	BI081911 602877135
40	343.8	27.9	723	10	BG971566 602840616
41	333	27.0	507	10	BE907720 601497752
42	331.8	26.9	650	9	AI451851 mx19h09.y
43	327.6	26.6	530	9	AI597102 m171f07.y
44	327.2	26.5	585	10	BG894625 355310 MA
45	318.6	25.8	1916	10	BF785809 602112433

ALIGNMENTS

RESULT 1
BI517905
LOCUS
DEFINITION
603041985F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5182556 5',
mRNA sequence.
BI517905
ACCESSION
BI517905.1 GI:15342697
VERSION
KEYWORDS
SOURCE
ORGANISM
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/.

REFERENCE
1 (bases 1 to 763)
AUTHORS
TITLE
JOURNAL
COMMENT
Unpublished (1999)
National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strassberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM11435 row: 1 column: 21
High quality sequence stop: 739.
Location/Qualifiers
1. 763
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5182556"
/clone_lib="NIH_MGC_116"
/lab_host="DH10B"
/note="Organ: pooled colon, kidney, stomach; Vector:
pCMV-SPORT6; Site:1: NotI; Site:2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons; age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
stomachs, 62 yo male and 70 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is

FEATURES

source

destroyed upon cloning). Average insert size 1.4 kb.
Insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is a NIH-MGC Library."

BASE COUNT 121 a 269 c 229 g 143 t 1 others

ORIGIN

Query Match 59.4%; Score 733; DB 10; Length 763;
Best local similarity 99.6%; Pred. No. 6,6e-121;
Matches 755; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 272 AAGTATATGGGGGCTTACCTGCTCCATCCAGAACATGAGCTTCCCTCCCTCACTCTC 331
DB 1 AAGTATATGGGGGCTTACCTGCTCCATCCAGAACATGAGCTTCCCTCCCTCACTCTC 60
QY 332 AAGAGACCTGGGCTTACCAACGACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 391
DB 61 AAGAGACCTGGGCTTACCAACGACGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
QY 392 CCTGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 451
DB 121 CCTGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
QY 452 AAGACGCGATATGGGAGAGTGAATAAACGACGAGACCTCTAGACGCTACGCTCTCT 511
DB 181 AAGACGCGATATGGGAGAGTGAATAAACGACGAGACCTCTAGACGCTACGCTCTCT 240
QY 512 ACAGGACCTGCGCGGACGACGAGTCTGTAAGCTTCACTTAAGCGGAGCTGAGG 571
DB 241 ACAGGACCTGCGCGGACGACGAGTCTGTAAGCTTCACTTAAGCGGAGCTGAGG 300
QY 572 GGGGTGGGGGCTATCAAGCTCTCTGAGAGGACGAGCTCTGCTGCGGCGCTGAGGCT 631
DB 301 GGGGTGGGGGCTATCAAGCTCTCTGAGAGGACGAGCTCTGCTGCGGCGCTGAGGCT 360
QY 632 CGGCGGACCTCTGTAAGCTGAGAGGCTGCGGAGCTGAGTGTGTTGCGAGG 691
DB 361 CGGCGGACCTCTGTAAGCTGAGAGGCTGCGGAGCTGAGTGTGTTGCGAGG 420
QY 692 CTTCTCTGAGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 751
DB 421 CTTCTCTGAGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 752 AAGTCAACCGGACGACCTCTGATCACTTCACTTCACTTCACTTCACTTCACTTCACT 811
DB 481 AAGTCAACCGGACGACCTCTGATCACTTCACTTCACTTCACTTCACTTCACTTCACT 540
QY 812 CGGCGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 871
DB 541 CGGCGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 872 GCTCCGTCGATCTCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 931
DB 601 GCTCCGTCGATCTCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 659
QY 932 TGGGATACGAGGCGGCTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 991
DB 660 TGGGATACGAGGCGGCTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 718
QY 992 TGGTGAATTTCTTGAAGGCGAGTCCCTGAGAGGCGGGGCT 1029
DB 719 TGGTGAATTTCTTGAAGGCGAGTCCCTGAGAGGCGGGGCT 756

RESULT 2

LOCUS

DEFINITION BG740350 941 bp mRNA linear EST 15-MAY-2001

ACCESSION BG740350 mRNA sequence.

VERSION BG740350.1 GI:14051003

KEYWORDS

EST,

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 941)
AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapb-remail.nih.gov
Tissue procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL), DNA
Sequencing by: Invitrogen, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LAM1D636 Row: 1 Column: 02
High quality sequence stop: 856.
Location/Qualifiers

FEATURES
source
1..941
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:479337"
/lab_host="DH10B (RT phage-resistant)"
/note="Organ: skin; Vector: pCMV-Sport6; Site: 1; NotI;
Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI-CGAP Library."

BASE COUNT 173 a 308 c 299 g 161 t

ORIGIN

Query Match 58.8%; Score 725.2; DB 10; Length 941;
Best local similarity 98.1%; Pred. No. 1.6e-119;
Matches 808; Conservative 0; Mismatches 8; Indels 8; Gaps 7;

QY 411 GCTCTATGTCAGAGTCCCTGCTCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 470
DB 1 GCTCTATGTCAGAGTCCCTGCTCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
QY 471 GGAGATTAACGAGCGGAAAGCTCTAGACGCTTCTCTACAGGACCTGCCCCAGGA 530
DB 61 GGAGATTAACGAGCGGAAAGCTCTAGACGCTTCTCTACAGGACCTGCCCCAGGA 120
QY 531 CCGCAAGTTCGTAAGCTTCACTTAAGCCGACGCTGAGAGCGGCGCTGAGGAGCT 590
DB 121 CCGCAAGTTCGTAAGCTTCACTTAAGCCGACGCTGAGAGCGGCGCTGAGGAGCT 180
QY 591 CTTCCGAGAGGACGCGACCTCTGCGGCGGCTGAGAGCGGCGGCGGCGGCGGCGGCG 650
DB 181 CTTCCGAGAGGACGCGACCTCTGCGGCGGCTGAGAGCGGCGGCGGCGGCGGCGG 240
QY 651 CCGTGAAGCGGCGGAGCGCTCACTGAGGCTGCTTGGAGGCGCTTCTGAGGCGGCGG 710
DB 241 CCGTGAAGCGGCGGAGCGCTCACTGAGGCTGCTTGGAGGCGCTTCTGAGGCGGCGG 300
QY 711 GTCGAGCCAGACCTTCCGAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 770
DB 301 GTCGAGCCAGACCTTCCGAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 771 CTTTCATCACTTGAAGGCGCAGAGGCGGACCCGCGACCCGCGGCTGCTGCTGCTG 830
DB 361 CTTTCATCACTTGAAGGCGCAGAGGCGGACCCGCGACCCGCGGCTGCTGCTGCTG 420
QY 831 CCAGACGCGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 890
DB 421 CCAGACGCGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 891 CGATTTTGGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 949
DB 481 CGATTTTGGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540

QY 950 AAGAGACCCCGACGAGCTGC-AGGACGACAAGACCCCATGCTGATCTTTCGAGGC 1008
|||||
DB 541 AAGGAGACCCCGACGAGCTGC-AGGACGACAAGACCCCATGCTGATCTTTCGAGGC 600
QY 1009 CGAGTCCCTGAGGGCGGCGCTGG-ACTCAGAGTGGACCCCGAGCCTTGAGGGGAGCT 1067
|||||
DB 601 CGAGTCCCTGAGGGCGGCGCTGG-AGTGGAGTGGACCCCGAGCCTTGAGGGGAGCT 660
QY 1068 GAGTTCGGGGGCGCTGTTTGGAGAGCCATCAGCTCCACCGCAGCAGTGGGGGCTTC 1127
|||||
DB 661 GGGTTCGGGGGCGCTGTTTGGAGAGCCATCAGCTCCACCGCAGCAGTGGGGGCTTC 720
QY 1128 GCTGGG-AGAGAGCCGAGCA--GCGAAGTGGAGCTCTCGATCTCGGCTCGCGAACTA 1184
|||||
DB 721 GCTGGGAGAGAGCCGAGCAAGTGGAGCTCTCGATCTCGGCTCGCGAACTA 780
QY 1185 CAGTCCCGGACAC-GACTTCTACTGCT-GGTGTCCAGGATGA 1226
|||||
DB 781 CAGTCCCGGACACAGTCTTCTACTGCTCGGAGTGTCCAGGATGA 824

RESULT 3
BE797163 821 bp mRNA linear EST 20-SEP-2000
LOCUS 60158717f1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3941561 5',
DEFINITION mRNA sequence.
ACCESSION BE797163
VERSION BE797163.1 GI:10218361
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 821)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov
Plate: LCM793 row: m column: 18
High quality sequence stop: 793.
Location/Qualifiers
1..821

FEATURES
Source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3941561"
/clone_lib="NIH_MGC_7"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site: 1; XhoI; Site: 2;
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 141 a 270 c 250 g 160 t

Query Match 55.3%; Score 681.4; DB 10; Length 821;
Best Local Similarity 94.4%; Pred. No. 9.9e-112;
Matches 728; Conservative 0; Mismatches 41; Indels 2; Gaps 2;
QY 206 GGGTCAGGCGCAACCTGTCAAGGTGCTGTGTCA-CTGTCTGGGGGCTCAACCTGACC 264
|||||

DB 14 GGGTCAGGCGCAA-CTGTCAAGGTGCTGTGTGTCAATGTGTCTGGGGGTCAACGTGACC 72
QY 265 AGCACTGAAGTCTATGGGGCGCTTCACTGCTCATACCGAATCATGAGTCTCCCTTC 324
|||||
DB 73 AGCACTGAAGTCTATGGGGCGCTTCACTGCTCATACCGAATCATGAGTCTCCCTTC 132
QY 325 ACTCTTCAGAGAGCTGAGCCCTACAGACCCAGTGGCTGGCTGGCTCCCTCCCTGCTC 384
|||||
DB 133 ACTCTTCAGAGAGCTGAGCCCTACAGACCCAGTGGCTGGCTGGCTCCCTCCCTGCTC 192
QY 385 CTGCTGGCCCTGCTGCTGGCCGCTGCTCTATGTCAAGTCCGCTCAAGTGTGCTC 444
|||||
DB 193 CTGCTGGCCCTGCTGCTGGCCGCTGCTCTATGTCAAGTCCGCTCAAGTGTGCTC 252
QY 445 TGGTACAGAGAGCGGTATGGGGAGTGGAGATTAACGACGGGAACCTCTACAGCCCTAC 504
|||||
DB 253 TGGTACAGAGAGCGGTATGGGGAGTGGAGATTAACGACGGGAACCTCTACAGCCCTAC 312
QY 505 GTCTCTCAGAGAGCTGCGCCGAGAGCCGCAAGTTCGTAACCTTCATCTAAAGCCGAG 564
|||||
DB 313 GTCTCTCAGAGAGCTGCGCCGAGAGCCGCAAGTTCGTAACCTTCATCTAAAGCCGAG 372
QY 565 CTGAGCGCGCTGCGGGCTTACAGCTCTTCTGAGACCGGAGCTCTGCGCGCTC 624
|||||
DB 373 CTGAGCGCGCTGCGGGCTTACAGCTCTTCTGAGACCGGAGCTCTGCGCGCTC 432
QY 625 GAGCCCTCGCGGAGCTCTTGGTGAACCTGAAGCGGCTGCGGAGCTCTATGCTGCTT 684
|||||
DB 433 GAGCCCTCGCGGAGCTCTTGGTGAACCTGAAGCGGCTGCGGAGCTCTATGCTGCTT 492
QY 685 TCGAGCGCTTCTGAGCGCGGCTGCTGAGACCCAGCTTCCGAGAGGCTGTGCGG 744
|||||
DB 493 TCGAGCGCTTCTGAGCGCGGCTGCTGAGACCCAGCTTCCGAGAGGCTGTGCGG 552
QY 745 CTGCTGAGCTTACCCGAGACCATCTTTCATCATCTTTCGAGAGGCGGAGCGGAGCC 804
|||||
DB 553 CTGCTGAGCTTACCCGAGACCATCTTTCATCATCTTTCGAGAGGCGGAGCGGAGCC 612
QY 805 GCGGACCGGGGCTGCGGCTGCTGCGGAGACCGGCTGCTGCTGCTG 864
|||||
DB 613 GCGGACCTGCGGCTGCGGCTGCTGCGGAGACCGGCTGCTGCTGCTG 672
QY 865 AGGCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 924
|||||
DB 673 AGGCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 732
QY 925 CGGAAGTTCGGTACAGCGCGGTGGAAGAGACCCCGAGAGCGAGCTGACG 975
|||||
DB 733 GGAAGTTCGGGAACAGCGCGGTGGAAGAGACCCCGAGAGCGAGCTGACG 783

RESULT 4
LOCUS B1521062 852 bp mRNA linear EST 29-AUG-2001
DEFINITION 603081652f1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5220901 5',
LOCUS B1521062
DEFINITION mRNA sequence.
ACCESSION B1521062
VERSION B1521062.1 GI:15345854
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 852)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LAM1155 row: 9 column: 14
 High quality sequence start: 3
 High quality sequence stop: 780.

FEATURES

source

1..852
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5220901"
 /clone_11b="NIH_MGC_120"
 /lab_host="DH10B"
 /note="Organ: pooled pancreas and spleen; Vector: pCMV-Sport6; Site: 1: NotI; Site: 2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1.2-5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH_MGC Library."
 Bases: 138 a 284 c 255 g 175 t
 ORIGIN

Query Match

Best Local Similarity 98.7%; Pred. No. 5.4e-108;
 Matches 697; Conservative 0; Mismatches 6; Indels 3; Gaps 3;

QY 1 ATGCGAGTGTCTGTATAGGCGCCCTGACTCTCTCCGCTGAGAGCAGTGTCTG 60
 Dd 148 ATGCGAGTGTCTGTATAGGCGCCCTGACTCTCTCCGCTGAGAGCAGTGTCTG 207
 QY 61 AGGCGTGTGGGCGAGCTGAGTGGCTGGAACCTGACAGGCTGGGAGTGTGGGCGC 120
 Dd 208 AGGCGTGTGGGCGAGCTGAGTGGCTGGAACCTGACAGGCTGGGAGTGTGGGCGC 267
 QY 121 CACTGTCTCTGCTTCACTGAGTGGCTGGAAGGCGGCTTCCATTGGGAATGGGCGC 180
 Dd 268 CACTGTCTCTGCTTCACTGAGTGGCTGGAAGGCGGCTTCCATTGGGAATGGGCGC 327
 QY 181 CACTGAGCTCTGAGAGTACTCTGGGTCAAGGCCAAGCTGTCAAGAGTGTGTCTC 240
 Dd 328 CACTGAGCTCTGAGAGTACTCTGGGTCAAGGCCAAGCTGTCAAGAGTGTGTCTC 387
 QY 241 AGTGTCTGGGGGTCAAGCTGACAGCACTGAAGTCTATGGGGCTTCACTGCTCATC 300
 Dd 388 AGTGTCTGGGGGTCAAGCTGACAGCACTGAAGTCTATGGGGCTTCACTGCTCATC 447
 QY 301 CAGAATCATGCTTCT 360
 Dd 448 CAGAATCATGCTTCT 507
 QY 361 GGGGTCTGGGCGCT 420
 Dd 508 GGGGTCTGGGCGCT 567
 QY 421 AAGTGTCT 480
 Dd 568 AAGTGTCT 627
 QY 481 GAGGGAAGCTCTGAGAGCT 540
 Dd 628 GAGGGAAGCTCTGAGAGCT 687
 QY 541 GTGAATCTATCTAAAGCGAGCTGGA - GGGGGTGGGGGTCTAAAGTCTCTCTCTG 599
 Dd 688 GTGAATCTATCTAAAGCGAGCTGGA - GGGGGTGGGGGTCTAAAGTCTCTCTCTG 747
 QY 600 CGAGCGGAGCTCTGCGCGCTGAGCGCTCTGCGCGAGCTCTGAGGAGCTCTGGA - GCC 658
 Dd 748 CGAGCG - GAGCTCTGCGCGCTGAGCGCTCTGCGCGAGCTCTGAGGAGCTCTGAGGCC 806

QY 659 GCTGCCGAGCGCTCATGCTGTGTCTTTGAGCGCCTTCTGAGCGG 704
 Dd 807 GCTGCCGAGCGCTCATGCTGTGTCTTTGAGCGCCTTCTGAGCGG 852

RESULT 5

BI821947

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 749)
 NIH-MGC http://imgc.nhl.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LAM11450 row: d column: 16
 High quality sequence stop: 739.

FEATURES

source

1..749
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5180511"
 /clone_11b="NIH_MGC_115"
 /lab_host="DH10B"
 /note="Organ: pooled brain, lung, testis; Vector: pCMV-Sport6; Site: 1: NotI; Site: 2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age 27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."
 BASE COUNT 127 a 262 c 218 g 142 t
 ORIGIN

Query Match

Best Local Similarity 99.5%; Pred. No. 3.1e-104;
 Matches 641; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 203 CCTGGGTCAAGGCCAAGCTGTCAAGTGTCTTGTCTGAGTGTCTGAGGCTCAACGTGA 262
 Dd 106 CCTGGGTCAAGGCCAAGCTGTCAAGTGTCTTGTCTGAGTGTCTGAGGCTCAACGTGA 165
 QY 263 CCAGCAGTGAAGTCTATGGGGCTTCACTGCTGCTCATCTCAGAACATCAGTCTCTCTCT 322
 Dd 166 CCAGCAGTGAAGTCTATGGGGCTTCACTGCTGCTCATCTCAGAACATCAGTCTCTCTCT 225
 QY 323 TCACCTTCAAGAGAGTGTGCTTCAAGCCAGCGGTGTGGGCTGAGGCTCTCTCTCTG 382
 Dd 226 TCACCTTCAAGAGAGTGTGCTTCAAGCCAGCGGTGTGGGCTGAGGCTCTCTCTCTG 285
 QY 383 TCCTGTCTGGCGCTGTGTGGCGCGCTGCTGTATGTGTCAGTGTGCTGCTCAAGTGTCTG 442
 Dd 286 TCCTGTCTGGCGCTGTGTGGCGCGCTGCTGTATGTGTCAGTGTGCTGCTCAAGTGTCTG 345
 QY 443 TCTGTACAGAGCGCGTATGGGAGGTGAGATTAACGAGCGGAAGCTCTACGAGCGCT 502


```
|||||
Db 346 TCTGTACACGACGCGTATGGGAGTGAGATAAAGACGGGAGCTCTACGACGCT 405
Oy 503 AGCTCTCCACAGCAGCTACCTCCCGAGAGACCGAAGTCTGTAACCTATCTAAACCCG 562
Db 406 AGCTCTCCACAGCAGCTACCTCCCGAGAGACCGAAGTCTGTAACCTATCTAAACCCG 465
Oy 563 AGCTGAGCGGCGTGGGGCTACAAAGCTCTCTCTGAGACGCGGAGCTCTGCGCGG 622
Db 466 AGCTGAGCGGCGTGGGGCTACAAAGCTCTCTCTGAGACGCGGAGCTCTGCGCGG 525
Oy 623 CTGAGCCCTCCGCGGACCTCTTGTGTAACCTGAGCGCTGCGGACCGCTATCTGTGTC 682
Db 526 CTGAGCCCTCCGCGGACCTCTTGTGTAACCTGAGCGCTGCGGACCGCTATCTGTGTC 585
Oy 683 TTTCGAGCGCTCTCTGAGACCGGGGCTGTGTGACGACCAAGCTCTCGGAGGGCTGTGTC 742
Db 586 TTTCGAGCGCTCTCTGAGACCGGGGCTGTGTGACGACCAAGCTCTCGGAGGGCTGTGTC 645
Oy 743 GGCTGTGAGCTCACCCGACAGACCATCTTCACTTTCAGAGGCGACGAGCGGAC 802
Db 646 GGCTGTGAGCTCACCCGACAGCATCTTCACTTTCAGAGGCGACGAGCGGAC 705
Oy 803 CCGCGACCCGCGGCTCCGCTGCTGCGGCGGACGCGGACCTG 846
Db 706 CCGCGACCCGCGGCTCCGCTGCTGCGGCGGACGCGGACCTG 749
```

```
RESULT 6      808 bp      mRNA      linear      EST 25-SEP-2001
B1769076      603057053F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5206380 5',
LOCUS          mRNA sequence.
DEFINITION     B1769076
VERSION        B1769076.1 GI:15760654
KEYWORDS       EST.
SOURCE         human.
ORGANISM       Homo sapiens
                Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
```

```
REFERENCE      1 (bases 1 to 808)
AUTHORS        NIH-MGC http://mgi.nci.nih.gov/.
TITLE          National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL        Unpublished (1999)
COMMENT        Contact: Robert Strausberg, Ph.D.
                Email: cgapbs-remail.nih.gov
```

```
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M1517 row: 1 column: 13
High quality sequence stop: 807.
```

FEATURES

source

```
1. 808
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5206380"
/clone_11b="NIH_MGC_122"
/lab_host="DH10B"
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 026. Note:
this is a NIH_MGC Library."
BASE COUNT    150 a      262 c      231 g      165 t
```

ORIGIN

Query Match 50.0%; Score 616; DB 10; Length 808;
Best Local Similarity 97.4%; Pred. No. 4, 2e-100;
Matches 637; Conservative 0; Mismatches 15; Indels 2; Gaps 1;

```
Oy 4 CCAGGTGTCTGTATAGAGCCCTGACTTCCTCTCCCGCTGAAAGACGAGTGTGAG 63
Db 155 CCAGGTGTCTGTATAGAGCCCTGACTTCCTCTCCCGCTGAAAGACGAGTGTGAG 214
Oy 64 CCGCTGTGGGACGCTCACTGCTGAACTGACGCGTGGGTAGCTCTGTGGGCCAC 123
Db 215 CCGCTGTGGGACGCTCACTGCTGAACTGACGCGTGGGTAGCTCTGTGGGCCAC 274
Oy 124 TGTCTCTGCTTCACTCACTGCTGAAAGAGGCGCTTCATTTGGGAATTTGGGGCCAC 183
Db 275 TGTCTCTGCTTCACTCACTGCTGAAAGAGGCGCTTCATTTGGGAATTTGGGGCCAC 334
Oy 184 TACAGCTCCACAGTACTCTGAGGTCAAGGCCAACCTGTCAAGAGGTGCTGTGTCAGT 243
Db 335 TACAGCTCCACAGTACTCTGAGGTCAAGGCCAACCTGTCAAGAGGTGCTGTGTCAGT 394
Oy 244 GTCTGTGGGCTCACTGACACGACTGAAGTCTATGGGCTTCACTGCTTCATTCAG 303
Db 395 GTCTGTGGGCTCACTGACACGACTGAAGTCTATGGGCTTCACTGCTTCATTCAG 454
Oy 304 AACATAGCTTCTCTCTCACTCTCAAGAGTGTGGCTTCAAGCCACGAGTGTGG 363
Db 455 AACATAGCTTCTCTCTCACTCTCAAGAGTGTGGCTTCAAGCCACGAGTGTGG 514
Oy 364 GTGCTGGCTTCTCTCTCTGCTGCTGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 423
Db 515 GTGCTGGCTTCTCTCTCTGCTGCTGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 574
Oy 424 TGGCTGTCAAGCTGCTGCTGTGTGACAGAGCGCTATGGGAGGTGAGATTAAGCAG 483
Db 575 TGGCTGTCAAGCTGCTGCTGTGTGACAGAGCGCTATGGGAGGTGAGATTAAGCAG 634
Oy 484 GGAACCTCTAGACGCTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 543
Db 635 GGAACCTCTAGACGCTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 694
Oy 544 AACTTCATCTAAAGCGGAGCTGAGAGCGGCTGCGGGCTGACAA--GCTCTCTGAGAG 601
Db 695 AACTTCATCTAAAGCGGAGCTGAGAGCGGCTGCGGGCTGACAAAGCTTCTTGTGAG 754
Oy 602 ACCGAGCTCTGCTGCGGCGCTGAGCCCTCGCGGACCTCTGTTGTAACCTGA 655
Db 755 ACCGAGCTCTGCTGCGGCGCTGAGCCCTCGCGGACCTCTGTTGTAACCTGA 808
```

```
RESULT 7      791 bp      mRNA      linear      EST 29-AUG-2001
B1518182/c    603041985F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5182556 3',
LOCUS          mRNA sequence.
DEFINITION     B1518182
VERSION        B1518182.1 GI:15342974
KEYWORDS       EST.
SOURCE         human.
ORGANISM       Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
```

```
REFERENCE      1 (bases 1 to 791)
AUTHORS        NIH-MGC http://mgi.nci.nih.gov/.
TITLE          National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL        Unpublished (1999)
COMMENT        Contact: Robert Strausberg, Ph.D.
                Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
```


OY 301 CAGACATCAGCTTCTCCTCCCTTACCTCTTGAGAGAGCGCCCTAACAGCCAGCGGT 360
 Db 466 CAGAACATCAGCTTCTCCTCCTTACTCTTACAGAGCTGGCCCTTACAGACCACTGGGT 525
 OY 361 GCGGTCTGGCCCTCCCTCCTTGCTCTGCTGGCCCTCTGCTGCTATATGTC 420
 Db 536 GCGGTCTGGCCCTCCCTCCTTGCTCTGCTGGCCCTCTGCTGCTATATGTC 585
 OY 421 AAGTCCCTCTCAACGTCTGCTGTGTATACAGACGCGTATGGGAGGTGAGATTAAC 480
 Db 586 AAGTCCCTCTCAACGTCTGCTGTGTATACAGACGCGTATGGGAGGTGAGATTAAC 645
 OY 481 GACGGGAAGCTTTACGACGCTTACGCTTCTTACACGACCTGCCCCGAGAGCCGCAATGC 540
 Db 646 GACGGGAAGCTTTACGACGCTTACGCTTCTTACACGACCTGCCCCGAGAGCCGCAATGC 705
 OY 541 GTGAACCTTCATCTTAAACCCGACGCTGGAGAGCGGCT - - CGGGCTTACAAAGCTCTTCTGG 598
 Db 706 GTGAACCTTCATCTTAAACCCGACGCTGGAGAGCGGCTTGGGGGCTTAACAAGCTTCTTGG 765
 OY 599 ACGACCGGACCTCTGCGCGGC 621
 Db 766 ACGACCGGACCTCTGCGCGGC 788

RESULT 9	
B1518544	
LOCUS	
DEFINITION	B1518544 941 bp mRNA linear EST-29-AUG-2001
	603061665671 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5210965 5'
	mRNA sequence.

ACCESSION	BI518544	
VERSION	BI518544.1	GI:15343336
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	

REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 941)
TITLE	NIH-MGC http://mgc.ncl.nih.gov/
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT	Unpublished. (1999) Contact: Robert Strausberg, Ph.D.

```

Plate: LLAM1529 row: 1 column: 18
High quality sequence stop: 897.
Location/Qualifiers
FEATURES
    source          1..941

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5210969"
/clone_1lb="NH_MGC_118"
/tissue_type="leukocyte"
/lab_host="DH10B"
/Note="Vector: PCMV-SF0RT6; Site_1: NotI; Site_2: EcoRV
(destroyed); RNA source leukocytes from anonymous pool of
non-activated adult donors. Library is oligo-dT primed
and directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
1.2-3.3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Introgen). Research Genetics tracking code 027. Note:
+1000 bp upstream of the start of the gene."

```

BASE COUNT
ORIGIN

Query Match

45.28; Score 556.8; DB 10; Length 941;

Best Local Similarity 97.48; Pred. No. 1,4e-89;
Matches 608; Conservative 0; Mismatches 12; Indels 4; Gaps 4

OY	1	ATGCCAGGTCGTGTGATAGAGGCGCCCTAGACTCTCTCTCCCGGCTGAAACACGAGTGCTG	60
Db	291	ATGCCAGGTCGTGTGATAGAGGCGCCCTAGACTCTCTCTCCCGGCTGAAACACGAGTGCTG	350
OY	61	AGGCGCTGCCTTGGGAGGCTCAGTGGCTCTGAACTGCAGGCTTGGGTACTCTCTGGGCC	120
Db	351	AGGCGCTGCCTTGGGAGGCTCAGTGGGTCTGAACCTGCAGCGCTTGGGTACTCTCTGGGCC	410
OY	121	CACCTGCTCCCTGSCCTTCCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC	180
Db	411	CACCTGCTCCCTGSCCTTCCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC	470
OY	181	CACCTACGCTCCACGAGTACTCTGGGGTCAAGGCCAACCTGTACAGAGGTGCTTGTGTCC	240
Db	471	CACCTACGCTCCACGAGTACTCTGGGGTCAAGGCCAACCTGTACAGAGGTGCTTGTGTCC	530
OY	241	AGTGTCTGGGGGGTCAACGTGACACAGCACTGAATCTATGGGACCTTCACTGTCTCATTC	300
Db	531	AGTGTCTGGGGGGTCAACGTGACACAGCACTGAATCTATGGGACCTTCACTGTCTCATTC	590
OY	301	CAGAACATCAGCTTCCCTCCCTCACTCTTCAGAGAGTGGGCCCTAACAGCCACGCGGCT	360
Db	591	CAGAACATCAGCTTCCCTCCCTCACTCTTCAGAGAGTGGGCCCTAACAGCCACGCGGCT	650
OY	361	GGGTCGTGGGCTCCCTCTCTGTGCTGTGAGCCCTGCTGAGCGCCCTGCTCTATGTC	420
Db	651	GGGTCGTGGGCTCCCTCTCTGTGCTGTGAGCCCTGCTGAGCGCCCTGCTCTATGTC	710
OY	421	AAGTCCGCTCTAACGTGCTGCTGTGTGTACACAGAGCGGTATGGGAGGTGGAGATTAAC	480
Db	711	AAGTCCGCTCTAACGTGCTGCTGTGTGTACACAGAGCGGTATGGGAGGTGGAGATTAAC	770
OY	481	GACGGGAAGCTCTACGAGCGCTACGCTGCC -TACAGCGACTGCCCGAGAGACCGCAAGTT	539
Db	771	GACGGGAAGCTCTACGAGCGCTACGCTCTCTTACACGACACTGCCCCGAGAGACCGGAGGTT	830
OY	540	GCTGAAC -TTCACTCTAAAGCGGACGCTGAGCGGCGTGGGGCT -ACAAGCTCTTCTG	597
Db	831	GCTGAACCTTCTCACTTAAAGCGGACGCTGAGCGGCGTGGGGCTTAAAGAAGCTCTTCTG	890
OY	598	GACGACCGGAGACCTCTGCGCGGC 621	
Db	891	GACGACCG -GACCTCTTGGGCGGC 913	

RESULT	10
B1752545	
LOCUS	1067 bp mRNA linear EST-25-SEP-2001
DEFINITION	6030219356P1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5192701 5' , mRNA sequence.
ACCESSION	B1752545
VERSION	B1752545.1 GI:15744123
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 1067)	NIH-MGC	http://mgc.ncbi.nlm.nih.gov/	National Institutes of Health, Mammalian Gene Collection (MGC)	Unpublished (1999)
	Contact:	Robert Strausberg, Ph.D.		

Email: cgabds-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>

Fri Nov 8 09:52:58 2002

us-09-598-443-1.rst

Page 8

Plate: L14M1481 row: P column: 14
High quality sequence stop: 622.
Location/Qualifiers

Source
1. .1067
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5192701"
/clone_1ib="NIH_MGC_114"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-Sport6; Site: 1; Nccl; male brains, age range 23-27 yo. Library is oligo-primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb. Insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: This is a NIH_MGC Library."
BASE COUNT 234 a 357 c 343 g 133 t
ORIGIN

Match 44.3% Score 546.8; DB 10; Length 1067;
c Local Similarity 96.4%; Pred. No. 8.7e-88; Mismatches 17; Indels 6; Gaps 5;
Matches 613; conservative 0;
603 CCGGACGCTCCGCGCCCTGAGCCCTCCGCGACCTCTTGGTGAACCTGAGCCG-T 661
4 CCGCGAGCGCCCTCCCGAGAGCCCTCCGCGACCTCTTGGTGAACCTGAGCCGCT 63
662 GCGGACGCTCCGCGCCCTGAGCCCTCCGCGACCTCTTGGTGAACCTGAGCCGCT 63
64 CCGGACGCTCCGCGCCCTGAGCCCTCCGCGACCTCTTGGTGAACCTGAGCCGCT 63
722 GCTTCGCGGAGGCGCTCTGCGCGCTGAGCTCAACCCGACACCATCTTATCACT 781
124 GCTTCGCGGAGGCGCTCTGCGCGCTGAGCTCAACCCGACACCATCTTATCACT 183
782 TCGAGGCGCAGAGCGCGGACCCCGCGCTCCGCGCTCTGCGCGCAGCACCCG 841
184 TCGAGGCGCAGAGCGCGGACCCCGCGCTCCGCGCTCTGCGCGCAGCACCCG 243
842 ACCGCTGACCTCTGCTGCTGAGGCGCGCTCCGCGCTCTGCGCGCAGCACCCG 901
244 ACCGCTGACCTCTGCTGCTGAGGCGCGCTCCGCGCTCTGCGCGCAGCACCCG 303
902 AAGAAGGAGCTGGC-GCTGCGCGGAGAGTGGCTTCAAGCCGCTGAGAGAGACCC 960
304 AAGAAGGAGCTGGC-GCTGCGCGGAGAGTGGCTTCAAGCCGCTGAGAGAGACCC 363
961 CAGACGACCTGAGAGCAGCAAGAGCCCATGCTGATTTCTTGAAGCCGCTGAG 1020
364 CAGACGACCTGAGAGCAGCAAGAGCCCATGCTGATTTCTTGAAGCCGCTGAG 423
1021 GCGCGGCGCTGAGCTCAAGAGTGGAGCCGCGACCTGAGGCGACCTGGTTCGGGG 1080
424 GCGCGGCGCTGAGCTCAAGAGTGGAGCCGCGACCTGAGGCGACCTGGTTCGGGG 483
1081 CCGTGTGAGAGCCCTGAGCTCAAGAGTGGAGCCGCGACCTGAGGCGACCTGGTTCGGGG 1139
484 CCGTGTGAGAGCCCTGAGCTCAAGAGTGGAGCCGCGACCTGAGGCGACCTGGTTCGGGG 543
1140 CCGGAG--CAAGGAGTGGAGCTGCGGATTCGCGCGGAGAACTACAGTGGCGGAC 1197
544 CCGGAGGAGGAGAAAGTGGAGCTGCGGATTCGCGCGGAGAACTACAGTGGCGGAC 603
1198 GACTTCTACTGCTGCTGCTCAAGGATGATGTAG 1233
604 GACTTCTACTGCTGCTGCTCAAGGATGATGTAG 638

RESULT 11
B1759065 983 bp mRNA linear EST 25-SEP-2001
LOCUS B1759065

DEFINITION 603042860F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5183271 5', mRNA sequence.
ACCESSION B1759065.1 GI:15750643
VERSION EST.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE NIH-MGC http://img.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Invitrogen, Inc.
Clone Distribution: MGC cDNA distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L14M1457 row: 9 column: 16
High quality sequence stop: 885.
Location/Qualifiers
1. .983
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5183271"
/clone_1ib="NIH_MGC_116"
/lab_host="DH10B"
/note="Organ: Pooled colon, kidney, stomach; Vector: pCMV-Sport6; Site: 1; Nccl; male brains, age range 23-27 yo. Library is oligo-primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb. Insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 023. Note: this is a NIH_MGC Library."
BASE COUNT 157 a 326 c 299 g 201 t
ORIGIN

Query Match 43.9% Score 541; DB 10; Length 983;
Best Local Similarity 91.0%; Pred. No. 9.3e-87;
Matches 711; conservative 0; Mismatches 25; Indels 45; Gaps 11;
1 ATGCCAGTGTCTGTGATAGGCGCCGCTGACTGCTCCGCTGAGAGACCTGCTG 60
217 ATGCCAGTGTCTGTGATAGGCGCCGCTGACTGCTCCGCTGAGAGACCTGCTG 276
61 AGGCTGCTGAGGCGAGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 120
277 AGGCTGCTGAGGCGAGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 336
121 CACTGCTCCCTGCTTCACTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 180
337 CACTGCTCCCTGCTTCACTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 396
181 CACTACAGCTCCAGAGTACTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 240
397 CACTACAGCTCCAGAGTACTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 456
241 AGTTCCTGAGGAGTCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 299
457 AGTTCCTGAGGAGTCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 516
300 CCAAGACATCACTTCTCTCTTCACTTCTTCAAGAGAGTGTGAGTGTGAGTGTGAGTGTG 359
517 CCAAGACATCACTTCTCTCTTCACTTCTTCAAGAGAGTGTGAGTGTGAGTGTGAGTGTG 576

[illegible]

Query Match	Similarity	Score	535.8	DB 10	Length 796	
Best Local	Similarity 97.1%	Pred. No. 7.6e-86				
Matches 609	Conservative 0	Mismatches 12	Indels 6	Gaps 6		
QY 1	ATGCCAGGTCGTGTGATAGAGGCGCCCTGACTTCCTCCCTCCCTGTAAGACAGGTGCTG	60				
Db 171	ATGCCAGGTCGTGTGATAGAGGCGCCCTGACTTCCTCCCTCCCTGTAAGACAGGTGCTG	230				
QY 61	AGGCGTCCTTGGGACGCTCACTGCTCTGAACTGCACGGCTTGGATGCTCTCGGCC	120				
Db 231	AGGCGTCCTTGGGACGCTCACTGCTCTGAACTGCACGGCTTGGATGCTCTCGGCC	290				
QY 121	CACGTGCTCCCTCCCTTTCAGTCCA -TGGCGTGAAGAACGGGGCTTCATGGGAAATGGGG	179				
Db 291	CACGTGCTCCCTCCCTTTCAGTCCATGTGGGTGAAGACGGGGCTTCATGGGAAATGGGG	350				
QY 180	CCACTACAGACCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGCAGAGTCTGTGTC	239				
Db 351	CCACTACAGACCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGCAGAGTCTGTGTA	410				
QY 240	CAGTGTCTGGGGGTCAACGTGACA -GCAGTGAAGTCTATGGGGCTTCACCTGCTCC -	297				
Db 411	CAGTGTCTGGGGGTCAACGTGACATGACATGGAAGTCTATGGGGCTTCACCTGCTCCG	470				
QY 298	ATCCAGAACATAGCTGTCCTCCCTTCACCTTCACCTTCAGAGAGTGGGCCCTTACAGCCAC	357				
Db 471	ATCCAGAACATAGCTGTCCTCCCTTCACCTTCACCTTCAGAGAGTGGGCCCTTACAGCCAC	529				
QY 358	GCTGGCGTGTGCTGGCCCTCCTGGTGTCTGCTGAGGCCCTGCTGAGCCGCCCTGCTCTAT	417				
Db 530	GCTGGCGTGTGCTGGCCCTCCTGGTGTCTGCTGAGGCCCTGCTGAGCCGCCCTGCTCTAT	589				
QY 418	GTCAGTGGCGTCTCAAGTGTGCTGTGCT -ACACAGACGCGTATGGGGAGTGGAGAT	476				
Db 590	GTCAGTGGCGTCTCAAGTGTGCTGTGCTGTGCTGTGACAGGACGCGTATGGGGAGTGGAGAT	649				
QY 477	AAAGCAGGGGAAGCTCTACGAGCGCTACGCTCTCTACACGAGCATGCCCGAGGACCG -CA	535				
Db 650	AAAGCAGGGGAAGCTCTACGAGCGCTACGCTCTCTCTACACGAGCATGCCCGAGGACCGCA	709				
QY 536	AGTTCGTGAACTTCAATCCCTAAAGCCGACGCTGAGCGGCGTCCGAGGCTTCTTC	595				
Db 710	AGTTCGTGAACTTCAATCCCTAAAGCCGACGCTGAGCGGCGTCCGAGGCTTCTTC	769				
QY 596	TGGACGACCGGACGCTCTGCGCGCG 622					
Db 770	TGGACGACCGGACGCTCTGCGCGCG 796					
RESULT 13						
LOCUS	Bi770099	801 bp	mRNA	linear	EST 25-SEP-2001	
DEFINITION	603053464P1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5203070 5',					
ACCESSION	Bi770099					
VERSION	Bi770099.1					
KEYWORDS	EST.					
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
TITLE	1 (bases 1 to 801)					
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/ .					
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)					
	unpublished (1999)					
	Contact: Robert Strausberg, Ph.D.					
	Email: rsrausb@mail.nih.gov					
	Tissue Procurement: Life Technologies, Inc.					
	cDNA Library Preparation: Life Technologies, Inc.					
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)					
	DNA Sequencing by: Incyte Genomics, Inc.					
	Clone distribution: MGC clone distribution information can be					

found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM1508 row: p column: 15
 High quality sequence stop: 801.
 Location/Qualifiers

FEATURES

source
 1. 801
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5203070"
 /clone_lib="NIH_MGC_122"
 /lab_host="DH10B"
 /note="Organ: pooled lung and spleen; Vector: PCMV-SPORT6;
 Site_1: NotI; Site_2: EcoRV (destroyed); RNA source:
 anonymous pool of 24 week female lung, 16 week female
 spleen, and 20-22 week male spleens. Library is digested
 primed, and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.4 kb, insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 026. Note:
 this is a NIH_MGC Library."
 COUNT 143 a 236 c 239 g 163 t

Query Match
 Best Local Similarity 42.4%; Score 522.4; DB 10; Length 801;
 Matches 545; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

OY 1 ATGCGAGTGTCTGTGATAGGCCCCCTGACTTCTCCCGCTGAAGACAGGTGCTG 60
 Db 254 ATGCGAGTGTCTGTGATAGGCCCCCTGACTTCTCCCGCTGAAGACAGGTGCTG 313
 OY 61 AGGCTGCTTGGGACCTAGTGGCTTGAACCTGACGCGCTTGGGTAGTCTCTGGGCC 120
 Db 314 AGGCTGCTTGGGACCTAGTGGCTTGAACCTGACGCGCTTGGGTAGTCTCTGGGCC 373
 OY 121 CACTGCTCCCTGCTTACGTCAGTGGCTGAAGACGGGCTTCCATTTGGCAATTGGGGCC 180
 Db 374 CACTGCTCCCTGCTTACGTCAGTGGCTGAAGACGGGCTTCCATTTGGCAATTGGGGCC 433
 OY 181 CACTAGACCTCCAGAGTACTCTGGGTCAAGGCCAACCTTGTCAAGAGTGTGTGTC 240
 Db 434 CACTAGACCTCCAGAGTACTCTGGGTCAAGGCCAACCTTGTCAAGAGTGTGTGTC 493
 OY 241 AGTGTCTGGGGGTCAAGGTGACGACGATGATGATGAGGCTTACCTGCTCCATC 300
 Db 494 AGTGTCTGGGGGTCAAGGTGACGACGATGATGATGAGGCTTACCTGCTCCATC 553
 OY 301 CAGAACATCAGCTTCTCCCTTCACTCTCAGAGAGCTGGCCCTTCAAGCCACAGTGGCT 360
 Db 554 CAGAACATCAGCTTCTCCCTTCACTCTCAGAGAGCTGGCCCTTCAAGCCACAGTGGCT 613
 OY 361 GCGGTGCTGGGCTCCCTCTCTGCTGTC -TGCCTGGCCCTGCTGTCGCGCCCTGCTATGT 419
 Db 614 GCGGTGCTGGGCTCCCTCTCTGCTGTC -TGCCTGGCCCTGCTGTCGCGCCCTGCTATGT 673
 OY 420 CAAAGTCCGTCAACGCTGCTGTGTATACAGAGCGCTA -TGGGAGGTGAGATTA 478
 Db 674 CAAAGTCCGTCAACGCTGCTGTGTATACAGAGCGCTA -TGGGAGGTGAGATTA 733
 OY 479 ACGAGGGAGACCTTACAGAGCCCTACAGCTCTCAGAGAGCTGCCCCGAGAGCCGCACT 538
 Db 734 ACGAGGGAGACCTTACAGAGCCCTACAGCTCTCAGAGAGCTGCCCCGAGAGCCGCACT 793
 OY 539 TCGTGAAC 546
 Db 794 TCGTGAAC 801

RESULT 14
 BIR20133 783 bp mRNA linear EST 04-OCT-2001
 LOCUS BIR20133
 DEFINITION 603037143F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5178226 5',

ACCESSION mRNA sequence.
 BIR20133
 VERSION BIR20133.1 GI:15931683
 KEYWORDS EST.
 SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 783)
 NIH-MGC http://mgs.nci.nih.gov/

Unpublished (1999)
 National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT

Journal
 Title
 Contact: Robert Strausberg, Ph.D.
 Email: ccgabs@email.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM1444 row: e column: 11
 High quality sequence stop: 701.
 Location/Qualifiers

FEATURES

source

1. 783
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5178226"
 /clone_lib="NIH_MGC_115"
 /lab_host="DH10B"
 /note="Organ: pooled brain, lung, testis; Vector:
 PCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
 source: anonymous pool of 6 male brains, age 69
 male lung, age 27; and 1 male testis, age 69
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.8 kb,
 insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 021. Note: this is a NIH_MGC Library."
 BASE COUNT 141 a 244 c 234 g 164 t

Query Match
 Best Local Similarity 41.6%; Score 513; DB 10; Length 783;
 Matches 599; Conservative 0; Mismatches 15; Indels 7; Gaps 7;

OY 1 ATGCGAGTGTCTGTGATAGGCCCCCTGACTTCTCCCGCTGAAGACAGGTGCTG 60
 Db 164 ATGCGAGTGTCTGTGATAGGCCCCCTGACTTCTCCCGCTGAAGACAGGTGCTG 223
 OY 61 AGGCTGCTTGGGACCTAGTGGCTTGAACCTGACGCGCTTGGGTAGTCTCTGGGCC 120
 Db 224 AGGCTGCTTGGGACCTAGTGGCTTGAACCTGACGCGCTTGGGTAGTCTCTGGGCC 283
 OY 121 CACTGCTCCCTGCTTACGTCAGTGGCTGAAGACGGGCTTCCATTTGGCAATTGGGGCC 180
 Db 284 CACTGCTCCCTGCTTACGTCAGTGGCTGAAGACGGGCTTCCATTTGGCAATTGGGGCC 343
 OY 181 CACTAGACCTCCAGAGTACTCTGGGTCAAGGCCAACCTTGTCAAGAGTGTGTGTC 239
 Db 344 CACTAGACCTCCAGAGTACTCTGGGTCAAGGCCAACCTTGTCAAGAGTGTGTGTC 403
 OY 240 CAGTGTCTGGGGGTCAACGTGACAGCACTGAAGTATAGGGGCTTCACTCTGTCAT 299
 Db 404 CAGTGTCTGGGGGTCAACGTGACAGCACTGAAGTATAGGGGCTTCACTCTGTCAT 462
 OY 300 CAGAACATCAGCTTCTCCCTTCACTCTCAGAGAGCTGGCCCTTCAAGCCAC -CTGG 358
 Db 463 CAGAACATCAGCTTCTCCCTTCACTCTCAGAGAGCTGGCCCTTCAAGCCAC -CTGG 522
 OY 359 CTGGGCTGCTGGCCCTCTCTGCTGCTGTCGCTTGTGTCGCGCCCTGCTCTATG 418

THIS PAGE BLANK (USPTO)